Erdős-Ko-Rado type problems, discrete isoperimetric inequalities and other problems in extremal combinatorics

by

William Raynaud

Submitted in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

School of Mathematical Sciences
Queen Mary, University of London
United Kingdom

April 2020

Declaration

I, William Raynaud, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 14/04/2020

Details of collaboration and publications:

Chapter 2 is based on work with Peter Cameron and my supervisor David Ellis which was published as "Smallest cyclically covering subspaces of $\mathbb{F}_q^n$ and lower bounds in Isbell's conjecture" in the European Journal of Combinatorics. We worked closely on the majority of the material in this chapter, though the work on symmetric covering subspaces of $\mathbb{F}_q^n$ is unpublished and is due to David Ellis and myself.

Chapter 3 is based on joint work with Cyrus Rashtchian which resulted in the preprint "Edge Isoperimetric Inequalities for Powers of the Hypercube" which we have submitted for publication.

Chapter 4 is joint work with my supervisor David Ellis. We worked closely together on the material of this chapter, and it is all joint work.

List of publications

Many of the ideas, chapters and sections of this thesis are based on manuscripts which are either published or in preparation, these are listed below.

- P. Cameron, D. Ellis, W. Raynaud, "Smallest cyclically covering subspaces of $\mathbb{F}_q^n$ and lower bounds in Isbell's conjecture," *European Journal of Combinatorics*, vol. 81, p. 242-255, Oct 2019.

- C. Rashtchian, W. Raynaud, "Edge Isoperimetric Inequalities for Powers of the Hypercube," arXiv preprint: arXiv:1909.10435.

- D. Ellis, W. Raynaud, "Families of sets that are pairwise close," in preparation.

Abstract

In this thesis, we investigate three problems in extremal combinatorics, using methods from combinatorics, representation theory, finite field theory and probabilistic combinatorics.

Firstly, for a prime power q and a positive integer n, we say a subspace U of $\mathbb{F}_q^n$ is *cyclically covering* if the union of the cyclic shifts of U is equal to $\mathbb{F}_q^n$. We investigate the problem of determining the minimum possible dimension of a cyclically covering subspace of $\mathbb{F}_q^n$. This is a natural generalisation of a problem posed in 1991 by Cameron. We prove several upper and lower bounds, and for each fixed q, we answer the question completely for infinitely many values of n (which take the form of certain geometric series). Our results imply lower bounds for a well-known conjecture of Isbell, and a generalisation thereof, supplementing lower bounds due to Spiga. We also consider the analogous problem for general representations of groups, and also provide some results for natural representations of the symmetric group S_n.

Second, for positive integers n and r, we let Q_n^r denote the rth power of the n-dimensional discrete hypercube graph (i.e., vertex set $\{0,1\}^n$ and edges between 0-1 vectors separated by Hamming distance at most r). We study edge isoperimetric inequalities for this graph. Harper, Bernstein, Lindsey and Hart proved a best-possible edge isoperimetric inequality for this graph in the case $r = 1$. For each $r \geq 2$, we obtain an edge isoperimetric inequality for Q_n^r; our inequality is tight up to a constant factor depending only upon r. Our techniques also give an edge isoperimetric inequality for the 'Kleitman-West graph' (the graph whose vertices are the k-element subsets of $\{1, 2, ..., n\}$, where two k-element sets are joined by an edge if they have symmetric difference of size 2); this inequality is tight up to a factor of $2 + o(1)$ for sets of size $\binom{n-s}{k-s}$, where $k = o(n)$ and $s \in \mathbb{N}$.

Finally, for positive integers n and d, we say sets $A, B \subseteq [n]$ are *d-close* if the minimum

cyclic distance between some element $a \in A$ and some element $b \in B$ is at most d. We say a set system $\mathcal{A} \subseteq \mathcal{P}([n])$ is d-close if every pair of sets $A, B \in \mathcal{A}$ is d-close, and a say a pair of set systems $\mathcal{A}, \mathcal{B}$ is *cross d-close* if every pair of sets $A \in \mathcal{A}, B \in \mathcal{B}$ is d-close. We investigate the maximum possible sizes of such set systems, particularly for each non-negative integer k, the maximum possible size of k-uniform d-close set systems (i.e., d-close set systems $\mathcal{A} \subseteq [n]^{(k)}$). This is a natural extension of the well-known result of Erdős, Ko and Rado, which corresponds to the case $d = 0$. We prove tight, stable bounds for k at most a constant fraction of n depending on d. To do so we employ the junta method, introduced to extremal combinatorics by Dinur and Friedgut, and initially applied to Erdős-Ko-Rado type problems by Keller and Lifshitz.

Acknowledgments

Firstly, I would like to thank my supervisor David Ellis for his guidance and the exciting mathematics he has enabled me explore. His invaluable support was an enormous help to me throughout my research.

I am also very grateful to the Professional Services staff at the School of Mathematical Sciences for creating and maintaining a great working environment. I want to thank the Academic staff and especially those in the Combinatorics Group, for their constant enuthusiasm and for introducing me to a wide range of mathematical ideas.

I must also thank the co-authors with whom I worked on the papers that make up this thesis. I am very grateful to Peter Cameron, Cyrus Rashtchian and, again, David Ellis. I am extremely proud to have worked closely with such extraordinary mathematicians; it was an enormous privilege.

I would also like to extend personal thanks to my friends who have been a huge support to me throughout my PhD: Lewis, Rachael, Jack, Liam, Ben, Natalie, Oliver, Kieran, Mayank, Shaoxiong, Diego and Ali. Most of all I would like to thank my parents Frances and Bernard, and my sister Isabel, as well as my extended family, for all their love.

This work was supported by a QMUL Principal's Postgraduate Research Studentship, and I am grateful to Queen Mary University for the financial support which made my PhD possible.

Table of Contents

List of Figures

Chapter 1

Introduction

In this thesis, we investigate three different problems of Extremal Combinatorics, i.e., questions of following form: given a set of objects, S, what is the maximum (or possibly minimum) possible size of a family of objects in S which satisfies a given property, P? We will also be interested in characterising families of maximum (or minimum) size, which we call 'extremal' families with property P, and whether or not this characterisation is in some sense stable (i.e., families that are nearly maximum in size are small alterations of an extremal family).

1.1 Smallest cyclically covering subspaces of $\mathbb{F}_q^n$

In Chapter 2, we consider, for a prime power q, an extremal problem on vector spaces over finite fields $\mathbb{F}_q$. Specifically, for $n \in \mathbb{N}$ we let $\{e_1, ..., e_n\}$ denote the standard basis for $\mathbb{F}_q^n$, then let $\sigma : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be the linear map defined by $\sigma\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_{i-1} e_i$, where addition/subtraction in the index is modulo n, i.e., σ is the *cyclic shift* operator which moves every entry one place clockwise. For a subspace $U \leq \mathbb{F}_q^n$ and a linear map $\alpha : \mathbb{F}_q^n \to \mathbb{F}_q^n$ we let $\alpha(U) = \{\alpha(x) : x \in U\}$. In particular, for $r \in \{0, 1, 2, ..., n-1\}$, $\sigma^r(U)$ is the subspace of $\mathbb{F}_q^n$ obtained by cyclically shifting the elements of U precisely r places clockwise. We call $\{\sigma^r(U) : 0 \leq r \leq n-1\}$ the family of *cyclic shifts* of U.

We say that a subspace $U \leq \mathbb{F}_q^n$ is *cyclically covering* if $\bigcup_{r=0}^{n-1} \sigma^r(U) = \mathbb{F}_q^n$. A natural extremal question is to ask: for prime power q and $n \in \mathbb{N}$, what is the smallest possible size of a cyclically covering subspace of $\mathbb{F}_q^n$? This is a natural extension of a question originally asked by Cameron [24, Problem 190] seeking the smallest possible cyclically covering subspace of $\{x \in \{0,1\}^n \mid \sum_{i=1}^n x_i \text{ is even}\}$ and asking whether the codimension of such a subspace tends to infinity as n tends to infinity over the odd integers. After the publication of our results, Cameron's original question was resolved by Aaronson, Groenland and Johnston [1], assuming a conjecture of Artin [46] on the existence of infinitely many primes p for which 2 is a primitive root. Artin's conjecture is widely believed to be true, and was shown by Hooley [46] to be a consequence of the Riemann hypothesis.

We define $h_q(n)$ to be the maximum possible codimension of a cyclically covering subspace of $\mathbb{F}_q^n$. Hence the smallest possible size of a cyclically covering subspace of $\mathbb{F}_q^n$ is, by definition, $q^{n-h_q(n)}$. We work with codimension as this turns out to be the more natural parameter for the problem. We prove several upper and lower bounds for $h_q(n)$. For each prime power q, the main results from Chapter 2 determine $h_q(n)$ completely for infinitely many values of n (which take the form of certain geometric series).

These results imply lower bounds for a conjecture of Isbell [48, 49], and a generalisation thereof [17], supplementing lower bounds due to Spiga [69]. To state Isbell's conjecture we require some notation. For $n \in \mathbb{N}$ we let $[n] := \{1, 2, ..., n\}$ and denote by S_n the symmetric group on $[n]$. A permutation group $G \leq S_n$ is *transitive* if for every $i, j \in [n]$ there exists a $\sigma \in G$ such that $\sigma(i) = j$. For a finite set X we define the *power set* of X to be $\mathcal{P}(X) := \{S \subseteq X\}$. A family of sets $\mathcal{F} \subseteq \mathcal{P}([n])$ is called *intersecting* if any two sets in $\mathcal{F}$ have non-empty intersection. A family $\mathcal{F}$ is an *up-set* if whenever $S \in \mathcal{F}$ and $S \subseteq T$ then $T \in \mathcal{F}$. We say that a family $\mathcal{F}$ is *antipodal* if for any $S \subseteq [n]$, $\mathcal{F}$ contains exactly one of S and $[n] \setminus S$. For $\mathcal{F} \subseteq \mathcal{P}([n])$ we define its *automorphism group* by $\mathrm{Aut}(\mathcal{F}) := \{\sigma \in S_n \ : \ \sigma(\mathcal{F}) = \mathcal{F}\}$, where $\sigma(\mathcal{F}) := \{\sigma(S) \ : \ S \in \mathcal{F}\}$, and we say that $\mathcal{F}$ is *symmetric* if $\mathrm{Aut}(\mathcal{F})$ is transitive. Isbell [48, 49], and later Cameron, Frankl and

Kantor [17] investigated the set

$$A := \{n \in \mathbb{N} \,:\, \text{there exists a symmetric intersecting family } \mathcal{F} \subseteq \mathcal{P}([n]) \text{ with } |\mathcal{F}| = 2^{n-1}\}.$$

It is easy to see that if $\mathcal{F} \subseteq \mathcal{P}([n])$ is symmetric, intersecting and $|\mathcal{F}| = 2^{n-1}$ then this is equivalent to $\mathcal{F}$ being an antipodal up-set. Symmetric, antipodal up-sets arise naturally as the 'winning sets' in n-player games where n different players are choosing between two alternatives and their choices are aggregated according to some rule. Symmetry is a a natural notion of the fairness of this aggregation rule. Isbell was led to study such set families from problems in Social Choice Theory. Isbell observed that the set A is equal to the set

$$A_2 := \left\{ n \in \mathbb{N} \,:\, \begin{array}{l} \text{there exists a transitive group } G \leq S_n \text{ with no} \\ \text{fixed-point-free element of 2-power order} \end{array} \right\},$$

and conjectured that there exists a function $m \,:\, \{b \in \mathbb{N} \,:\, b \text{ odd}\} \to \mathbb{N}$ such that if $b \in \mathbb{N}$ is odd and $a \geq m(b)$, then $2^a \cdot b \notin A_2$. This conjecture, and a generalisation thereof [17], remains open. Our results provide lower bounds on the function m.

In addition to the problem of finding smallest possible cyclically covering subspaces, we also investigate the analogous problem for general representations of groups, and provide some additional results for the specific case of representations of the symmetric group S_n. The proofs use arguments from combinatorics, representation theory and finite field theory.

1.2 Edge isoperimetric inequalities for powers of the hypercube

In Chapter 3, we prove several *isoperimetric inequalities*: an isoperimetric inequality provides lower bounds on the minimum possible 'boundary size' among sets of a given 'size', where the precise meaning of these terms varies according to the problem. A

classical isoperimetric problem asked for the minimum possible perimeter among shapes in the plane that have area 1. The fact that it is best to take a circle was 'known' to the ancient Greeks, but was not proven rigorously until the 19th century when a proof was given by Weierstrass in a series of lectures in 1870's Berlin.

This classical isoperimetric problem has been solved with the ambient space replaced with n-dimensional Euclidean space, with the n-dimensional unit sphere, and with n-dimensional hyperbolic space with the natural notion of boundary in each case being surface area for sufficiently 'nice' sets. For background on isoperimetric inequalities we refer the reader to the survery of Osserman [65].

During the last half century, *discrete isoperimetric problems* have been extensively studied. These concern notions of boundary in graphs. There are two common notions of boundary in graph theory: for a fixed graph $G = (V, E)$ we have

- *Vertex boundary:* For a subset $A \subseteq V$, the vertex boundary is defined to be

$$\partial_v(A) := \{u \in V \setminus A \ : \ uv \in E \text{ for some } v \in A\}.$$

  The *vertex-isoperimetric problem for* G asks for the minimum possible size of the vertex-boundary of a m-element subset of V, for each $m \in \mathbb{N}$.

- *Edge boundary:* For a subset $A \subseteq V$, the edge boundary is defined to be

$$\partial(A) := \{uv \in E \ : \ v \in A, u \in V \setminus A\}.$$

  The *edge-isoperimetric problem for* G asks for the minimum possible size of the edge-boundary of a m-element subset of V, for each $m \in \mathbb{N}$.

Chapter 3 concerns edge-isoperimetric inequalities, i.e., lower bounds on

$$\min\{|\partial(A)| : A \subseteq V(G), |A| = m\}$$

for each integer m. For $n, r \in \mathbb{N}$ we let Q_n^r denote the rth power of the n-dimensional discrete hypercube graph, i.e., the graph with vertex-set $\{0,1\}^n$, where two $0-1$ vectors are joined if they are Hamming distance at most r apart. We prove several edge-isoperimetric inequalities for this graph. A best-possible edge isoperimetric inequality in the case $r = 1$ was proved by Harper [38], Bernstein [11], Lindsey [59] and Hart [42]. For each $r \geq 2$, we obtain an edge isoperimetric inequality for Q_n^r that is tight up to a constant factor depending only on r. These results also give an edge isoperimetric inequality for the 'Kleitman-West graph' [40], i.e., for an integer $0 \leq k \leq n$ the graph with vertex set $[n]^{(k)} := \{A \subset [n] : |A| = k\}$, and sets $A, B \in [n]^{(k)}$ joined if they have symmetric difference of size two.

Roughly simultaneously to our own results, Kirshner and Samorodnitsky [55] independently obtained similar results, using very different methods. The isoperimetric inequalities that can be deduced from Kirshner and Samorodnitsky's results are in some cases weaker than our own and in other cases stronger.

These isoperimetric problems have many applications, not only to other problems within mathematics but areas such as distributed algorithms [2, 10], communication complexity [38], network science [12] and game theory [42].

1.3 Families of sets that are pairwise close

In Chapter 4 we consider an extremal problem that is a natural extension of the classical Erdős-Ko-Rado theorem [28] for *intersecting families* of sets: for $n \in \mathbb{N}$ a family $\mathcal{A} \subset \mathcal{P}([n])$ is *intersecting* if for every pair of sets $A, B \in \mathcal{A}$ the intersection $A \cap B$ is non-empty. The Erdős-Ko-Rado theorem states that for $n, k \in \mathbb{N}$ such that $2k \leq n$, if $\mathcal{A} \subseteq [n]^{(k)}$ is intersecting (i.e., $\mathcal{A}$ is a *k-uniform intersecting family*), then $|\mathcal{A}| \leq \binom{n-1}{k-1}$. Furthermore, if $2k < n$ then the unique extremal intersecting families are isomorphic to $\{A \in [n]^{(k)} : 1 \in A\}$, and these families show the bound is tight for all k and n.

The Erdős-Ko-Rado theorem has been extended in numerous ways. The complete

intersection theorem of Ahlswede and Khachatrian [6, 7] extends the notion of intersecting to *t-intersecting* for each $t \in \mathbb{N}$: a set system $\mathcal{A} \subseteq \mathcal{P}([n])$ is t-intersecting if for each $A, B \in \mathcal{A}$ the intersection $A \cap B$ has size at least t. There have also been extensions which change the groundset $[n]$ such as Ellis, Filmus and Friedgut's [25] upper bound for triangle-intersecting families of graphs: a set of graphs $\mathcal{G}$ on vertex set $[n]$ is triangle intersecting if for each $G, H \in \mathcal{G}$ the intersection $G \cap H$ contains a triangle. Another example is a result of Ellis, Friedgut and Pilpel [26] which proves an Erdős-Ko-Rado type theorem for k-intersecting families of permutations: a subset $I \subseteq S_n$ is k-intersecting if every pair $\sigma, \pi \in I$ satisfies $|\{i \in [n] : \sigma(i) = \pi(i)\}| \geq k$. Still further extensions add extra structure on the groundset such as the result of Talbot [71] which solves the Erdős-Ko-Rado problem for set systems of separated sets: a subset A of $[n]$ is separated if for all $a, b \in A$ the cyclic distance between a and b is at least 1.

For our extension we weaken the notion of intersecting pairs of sets in the following way: for $n \in \mathbb{N}$ we define *cyclic distance* on $[n]$ by

$$\text{dist} \ : \ [n] \times [n] \to \mathbb{Z}_{\geq 0}; \quad \text{dist}(a, b) := \min_{z \in \mathbb{Z}} \{|z| \ : \ b \equiv a + z \mod n\}.$$

Then, for n, d positive integers, we say a pair of sets $A, B \subseteq [n]$ are *d-close* if

$$\min_{a \in A, b \in B} \text{dist}(a, b) \leq d.$$

A set system $\mathcal{A} \subseteq \mathcal{P}([n])$ is *d-close* if every pair of sets $A, B \in \mathcal{A}$ is d-close, so in particular a 0-close set system is an intersecting family. Natural extremal questions are then: what is the maximum size of a k-uniform, d-close set system? What is the structure of a maximum k-uniform, d-close set system? These truly extend the Erdős-Ko-Rado theorem, which corresponds to the case $d = 0$.

We prove tight, stable bounds when k is at most a certain constant fraction of n depending on d, and determine the extremal families in this regime. The proof follows the *junta method* introduced to Extremal Combinatorics by Dinur and Friedgut [23],

who went on to apply the method to Erdős-Ko-Rado type problems. These applications were significantly extended by Keller and Lifshitz [54].

1.4 Notation

Here we record some of the notation that will be used throughout the thesis. We note also that some notation varies between the chapters, since each deals with quite different mathematics. However, each chapter is written to be self contained, and notation necessary to a particular chapter is defined within the chapter itself.

1.4.1 Set systems

For a finite set S, we let $|S|$ be the number of elements in S. We let $\mathcal{P}(S) = \{A : A \subseteq S\}$ denote the *power set* of S, and for a non-negative integer k we let $S^{(k)} = \binom{S}{k} := \{A \subseteq S : |A| = k\}$. We will call subsets $\mathcal{A} \subseteq \mathcal{P}(S)$ *set systems* or *set families*, and we call subsets $\mathcal{A} \subseteq S^{(k)}$ *k-uniform* set systems.

For each $n \in \mathbb{N}$, we let $[n] := \{1, 2, ..., n\}$.

1.4.2 Asymptotic 'big-O' notation

We will use asymptotic notation as we define below. Suppose that $X \subseteq \mathbb{R}$ contains a sequence tending to infinity. Let $f, g : X \to \mathbb{R}$. Then

- $f = O(g)$ if there exists $x_* \in \mathbb{R}$ and $0 < C \in \mathbb{R}$ such that for all $x \geq x_*$ we have $|f(x)| \leq C|g(x)|$,
- $f = o(g)$ if for all $\varepsilon > 0$ there exists $x_* \in \mathbb{R}$ such that for all $x \geq x_*$ we have $|f(x)| \leq \varepsilon|g(x)|$,
- $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$.

1.4.3 Subspace relation

For vector spaces V, W, both over the field $\mathbb{F}$, such that $W \subseteq V$ and for which addition and scalar multiplication in W are the restrictions of addition and scalar multiplication in V, we say that W is a *subspace* of V. This is denoted $W \leq V$.

Chapter 2

Smallest cyclically covering subspaces of $\mathbb{F}_q^n$

2.1 Introduction

For a prime power q, let $\mathbb{F}_q$ denote the finite field of order q. For $n \in \mathbb{N}$, let $\{e_1, e_2, ..., e_n\}$ denote the standard basis for $\mathbb{F}_q^n$. Let $\sigma : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be the linear map defined by $\sigma(\sum_{i=1}^n x_i e_i) = \sum_{i=1}^n x_{i-1} e_i$, where addition/subtraction in the index is modulo n. That is, σ is the *cyclic shift* operator which shifts each entry one place clockwise. Given a subspace $U \leq \mathbb{F}_q^n$ and a linear map $\alpha : \mathbb{F}_q^n \to \mathbb{F}_q^n$ we let $\alpha(U) = \{\alpha(x) : x \in U\}$. In particular, for $r \in \{0, 1, 2, 3, ..., n-1\}$, $\sigma^r(U)$ is the subspace of $\mathbb{F}_q^n$ obtained by cyclically shifting the elements of U precisely r places clockwise. We call $\{\sigma^r(U) : 0 \leq r \leq n-1\}$ the family of *cyclic shifts* of U.

We say a subspace $U \leq \mathbb{F}_q^n$ is *cyclically covering* if $\bigcup_{r=0}^{n-1} \sigma^r(U) = \mathbb{F}_q^n$. For a prime power q and $n \in \mathbb{N}$, we define $h_q(n)$ to be the maximum possible codimension of a cyclically covering subspace of $\mathbb{F}_q^n$. The main purpose of this chapter is to investigate the behaviour of the function $h_q : \mathbb{N} \to \mathbb{N}$, for various prime powers q.

This problem is a natural generalisation of the following problem, posed by Cameron

in [24, Problem 190]. For $n \in \mathbb{N}$, define $V_n = \{x \in \mathbb{F}_2^n : \sum_{i=1}^n x_i = 0\}$, i.e., V_n is the $\mathbb{F}_2$-vector space of binary strings with length n and even Hamming weight. For an odd positive integer n, define $f(n)$ to be the maximum possible codimension of a subspace $W \leq V_n$ such that the union of the cyclic shifts of W is equal to V_n. Cameron asked whether $f(n)$ tends to infinity as $n \to \infty$ (over odd integers n).

We observe that $f(n) = h_2(n)$ for all odd $n \in \mathbb{N}$. Indeed, take $W \leq V_n$ such that V_n is equal to the union of the cyclic shifts of W. Then $W' := \mathrm{Span}(W \cup \{11...1\})$ is a cyclically covering subspace of $\mathbb{F}_2^n$ with the same codimension as that of W in V_n. Conversely, if $U \leq \mathbb{F}_2^n$ is a cyclically covering subspace, then the cyclic shifts of $U' := U \cap V_n$ cover V_n, and the codimension of U' in V_n is equal to the codimension of U in $\mathbb{F}_2^n$.

We remark that somewhat similar problems have been investigated before. In [61], for example, Luh shows that any vector space (finite or infinite) over $\mathbb{F}_q$ can be expressed as a union of $q+1$ proper subspaces, and that this expression is unique up to automorphisms of the vector space. In [50], Jamison determined, for each $0 < k < n$, the minimum number of k-flats that are required to cover $\mathbb{F}_q^n \setminus \{0\}$. (Here, a *$k$-flat* is a translate of a k-dimensional subspace.)

2.1.1 Connections to Isbell's conjecture

Our results have implications for a well-known conjecture of Isbell (and a generalisation thereof), as we now describe. This was Cameron's original motivation for studying the behaviour of $f(n)$ for large odd values of n, though we also believe that this problem (and our generalisation) is natural in its own right.

For $n \in \mathbb{N}$, we let $[n] := \{1, 2, \ldots, n\}$ denote the standard n-element set, and we write S_n for the symmetric group on $[n]$. If $G \leq S_n$ is a permutation group, we say that n is the *degree* of G, and we say that G is *transitive* if for every $i, j \in [n]$, there exists $\sigma \in G$ such that $\sigma(i) = j$. If X is a finite set, we write $\mathcal{P}(X)$ for the power-set of X. If $\mathcal{F} \subset \mathcal{P}([n])$, we say $\mathcal{F}$ is *intersecting* if any two sets in $\mathcal{F}$ have nonempty intersection, we say it is an *up-set* if whenever $S \in \mathcal{F}$ and $S \subset T$ we have $T \in \mathcal{F}$, and

we say $\mathcal{F}$ is *antipodal* if for any $S \subset [n]$, $\mathcal{F}$ contains exactly one of S and $[n] \setminus S$. If $\mathcal{F} \subset \mathcal{P}([n])$, we define its *automorphism group* by $\mathrm{Aut}(\mathcal{F}) := \{\sigma \in S_n : \ \sigma(\mathcal{F}) = \mathcal{F}\}$, where $\sigma(\mathcal{F}) := \{\sigma(S) : \ S \in \mathcal{F}\}$, and we say that $\mathcal{F}$ is *symmetric* if $\mathrm{Aut}(\mathcal{F})$ is a transitive subgroup of S_n.

Isbell [48], and later Cameron, Frankl and Kantor [17], investigated the set

$$A := \{n \in \mathbb{N} : \text{there exists a symmetric intersecting family } \mathcal{F} \subset \mathcal{P}([n]) \text{ with } |\mathcal{F}| = 2^{n-1}\}. \tag{2.1}$$

Since for any $S \subset [n]$, an intersecting family $\mathcal{F} \subset \mathcal{P}([n])$ contains at most one of S and $[n] \setminus S$, we have $|\mathcal{F}| \leq 2^{n-1}$ for any intersecting $\mathcal{F}$. It is easy to see that an intersecting family $\mathcal{F} \subset \mathcal{P}([n])$ is maximal intersecting if and only if $|\mathcal{F}| = 2^{n-1}$, and that a family $\mathcal{F} \subset \mathcal{P}([n])$ is maximal intersecting if and only if it is an antipodal up-set. Symmetric, antipodal up-sets arise naturally as the 'winning sets' in n-player games where n different players are choosing between two alternatives and their choices are aggregated according to some rule (with symmetry being a natural notion of the fairness of the rule), and Isbell [48] was led to their study from problems in Social Choice Theory. (Indeed, Isbell termed a symmetric, antipodal up-set a *fair game*, though we do not use this terminology here, to avoid confusion with other notions of fair games.) Isbell [49] observed that the set A defined in (2.1) is equal to the set of all positive integers n for which there exists a transitive permutation group of degree n having no fixed-point-free element of 2-power order. We provide a proof here for completeness. Indeed, we let

$$\begin{aligned} A_2 := \{n \in \mathbb{N} : &\text{ there exists a transitive permutation group of degree } n \\ &\text{with no fixed-point-free element of 2-power order}\}, \end{aligned}$$

and prove the following claim.

Claim 2.1.1. $A = A_2$

Proof. We show first that $A \subseteq A_2$. For each $n \in A$ there exists a symmetric intersecting

family $\mathcal{F} \subset \mathcal{P}([n])$ of size $|\mathcal{F}| = 2^{n-1}$, and so $\mathrm{Aut}(\mathcal{F}) \leq S_n$ is a transitive permutation group of degree n. Suppose that $\pi \in \mathrm{Aut}(\mathcal{F})$ is a fixed-point-free element of 2-power order. Then we can express π as product of disjoint cycles which partition $[n]$, i.e.,

$$\pi = (\pi_{11}\ \pi_{12}\ ...\ \pi_{1k_1})(\pi_{21}\ \pi_{22}\ ...\ \pi_{2k_2})...(\pi_{r1}\ \pi_{r2}\ ...\ \pi_{rk_r}),$$

and note the order of π is $\mathrm{lcm}(k_1, k_2, ..., k_r)$. Since π has 2-power order it follows that each k_i is a power of 2, and since π is fixed-point-free it follows that each $k_i \geq 2$. We let

$$X = \{\pi_{ij} : i = 1, 2, ..., r \text{ and } 1 \leq j \leq k_i \text{ is odd}\} \subset [n],$$

and note

$$\pi(X) = \{\pi_{ij} : i = 1, 2, ..., r \text{ and } 1 \leq j \leq k_i \text{ is even}\} = [n] \setminus X.$$

But since $\mathcal{F}$ is antipodal, precisely one of X and $[n] \setminus X$ is an element of $\mathcal{F}$. It follows that $\mathcal{F}$ is not closed under the action of π, contradicting the fact that $\pi \in \mathrm{Aut}(\mathcal{F})$. Hence $\mathrm{Aut}(\mathcal{F})$ has no fixed-point-free element of 2-power order. Thus $n \in A_2$ and furthermore $A \subseteq A_2$.

Next we show $A_2 \subseteq A$. Suppose that $n \in A_2$, so there exists a transitive group $G \leq S_n$ with no fixed-point-free element of 2-power order. We use G to define a process which constructs a symmetric intersecting family in $\mathcal{P}([n])$ of size 2^{n-1}. We initialise the process with

$$\mathcal{F}_0 = \{X \subseteq [n] : |X| > n/2\},$$

which is evidently a symmetric intersecting family, with $G \leq \mathrm{Aut}(\mathcal{F}_0) = S_n$ and size at most 2^{n-1}. Now suppose we have constructed symmetric intersecting families $\mathcal{F}_0 \subset \mathcal{F}_1 \subset ... \subset \mathcal{F}_i$ such that $\mathcal{F}_i \subset \bigcup_{r \geq n/2} [n]^{(r)}$ and $G \leq \mathrm{Aut}(\mathcal{F}_i)$. If $\mathcal{F}_i$ is of size 2^{n-1} then we are done and we halt the process. So without loss of generality $|\mathcal{F}_i| < 2^{n-1}$, and

therefore there exists $X \in [n]^{(n/2)}$ such that $X, X^c \notin \mathcal{F}_i$. Let

$$\mathcal{F}_{i+1} = \mathcal{F}_i \cup \{g(X) : g \in G\},$$

so $\mathcal{F}_i \subset \mathcal{F}_{i+1}$, $\mathcal{F}_{i+1} \subset \bigcup_{r \geq n/2} [n]^{(r)}$ and $G \leq \mathrm{Aut}(\mathcal{F}_{i+1})$. Furthermore, $\mathcal{F}_{i+1}$ is intersecting since if $U, V \in \mathcal{F}_{i+1}$ then we have the following cases:

- $U, V \in \mathcal{F}_i$,
- $U \in \mathcal{F}_i, V \in \{g(X) : g \in G\}$,
- $U, V \in \{g(X) : g \in G\}$.

The first case is trivial as $\mathcal{F}_i$ is intersecting. For the second case, take $g \in G$ such that $V = g(X)$ and suppose that $U \cap g(X) = \emptyset$. Then, as $|U|, |g(X)| \geq n/2$ we must have $U = g(X)^c = g(X^c)$ which implies that $X^c = g^{-1}(U) \in \mathcal{F}_i$, contradicting the choice of X. Hence $U \cap V \neq \emptyset$.

For the final case, let $g, h \in G$ such that $V = g(X), U = h(X)$, and suppose that $g(X) \cap h(X) = \emptyset$. Then, as in the second case, we see $h(X) = g(X)^c = g(X^c)$ and so $X^c = g^{-1}h(X)$. We let $\pi = g^{-1}h \in G$ and note that π is fixed-point-free and has even order. Let $d \geq 1$ be an integer and a be an odd integer such that the order of π is $2^d \cdot a$. Then $\pi^a \in G$ has order 2^d and, since $\pi^a(X) = X^c$, it is fixed-point-free. This contradicts the definition of G, and so $U \cap V \neq \emptyset$. This completes the proof that $\mathcal{F}_{i+1}$ is intersecting.

This process must halt as $2^{n-1} \geq |\mathcal{F}_{i+1}| \geq |\mathcal{F}_i| + 1$, and at termination we have $\mathcal{F}_i$ a symmetric intersecting family in $\mathcal{P}([n])$ of size 2^{n-1}. Thus $n \in A$, and furthermore $A_2 \subseteq A$. This completes the proof that $A = A_2$. □

Isbell made the following conjecture about the set A.

Conjecture 2.1.1. *There exists a function $m : \{b \in \mathbb{N} : b \text{ odd}\} \to \mathbb{N}$ such that if $b \in \mathbb{N}$ is odd and $a \geq m(b)$, then $2^a \cdot b \notin A$.*

Isbell's conjecture remains open. Cameron, Frankl and Kantor [17] proved that for $b \in \{1, 3\}$ one can take $m(1) = 1$ and $m(3) = 2$ (which is best possible). They also made the following generalisation of Isbell's conjecture.

Conjecture 2.1.2. *For each prime p, we define A_p to be the set of all positive integers n for which there exists a transitive permutation group of degree n having no fixed-point-free element of p-power order. For each prime p, there exists a function $m_p : \{b \in \mathbb{N} : gcd(b, p) = 1\} \to \mathbb{N}$ such that if $b \in \mathbb{N}$ is coprime to p and $a \geq m_p(b)$, then $p^a \cdot b \notin A_p$.*

Hereafter, we refer to this as the *generalised Isbell conjecture*. For all primes p, the p-case of this conjecture remains open, although several related results have been proved; for example, using the Classification of Finite Simple Groups, Fein, Kantor and Schacher [29] proved that any transitive permutation group of degree $n > 1$ contains a fixed-point-free element of p-power order for some prime p.

It is natural to ask for lower bounds on the function m_p in the generalised Isbell conjecture. To obtain such a lower bound, it suffices to construct a transitive permutation group of degree $n = p^a \cdot b$, with b coprime to p, containing no fixed-point-free element of p-power order, and with a large compared to b. One construction method is to take the vector space $V = \mathbb{F}_p^b$, with b coprime to p, and to take a subspace $W \leq V$ of smallest possible dimension such that the cyclic shifts of W (i.e., the images of W under powers of the cyclic shift operator σ, defined above) cover V. Let $G := V \rtimes C_b$ be the semidirect product of V by the cyclic group $C_b = \langle\sigma\rangle$, and let $\iota : V \hookrightarrow G;\ v \mapsto (v, \mathrm{Id})$ denote the natural inclusion map. Consider the permutation group H induced by the left action of G on the left cosets of $\iota(W)$ in G. It is easy to see that every element of p-power order in H is induced by an element of G of the form (v, Id) for some $v \in V$; such an element fixes some left coset of $\iota(W)$ in G, namely $(0, \sigma^j)(\iota(W))$, where $j \in \{0, 1, \ldots, b-1\}$ is such that $v \in \sigma^j(W)$. Hence, any element of H of p-power order has a fixed point. The degree of H is $p^a \cdot b$, where $a := \dim(V) - \dim(W) = h_p(b)$ is the codimension of W. This yields the following.

Proposition 2.1.1. *For prime p and each $b \in \mathbb{N}$ coprime to p, define $m_p(b) = \min\{c \in \mathbb{N} : p^a \cdot b \notin A_p \ \forall a \geq c\}$, with the usual convention that $m_p(b) = \infty$ if the set in question is empty. Then $m_p(b) > h_p(b)$ for all integers $b \in \mathbb{N}$ that are coprime to p.*

This method led Cameron to the problem considered in this chapter. As we outline later, a construction due to Spiga [69] (building on the work of Suzuki in [70]) gives a lower bound for $m_p(b)$ which is better than ours for certain integers b; on the other hand, our construction is simpler and works for certain natural infinite sequences of integers where Spiga's method does not apply.

The rest of this chapter is structured as follows. In Section 2.2, we prove some simple upper and lower bounds on the function $h_q(n)$. In Section 2.3, we prove our main results, which are lower bounds on $h_q(n)$ that are tight for infinitely many values of n. In Section 2.4, we generalise the problem considered here to arbitrary group representations, and we prove some straightforward upper and lower bounds for the general problem. In Section 2.5, we exhibit, for each prime power q, infinitely many values of n for which $h_q(n) = 0$. We obtain this result in two ways: first as a special case of a result for arbitrary group representations where the only covering subspace is the whole space, and second from a direct combinatorial argument. Finally in Section 2.6 we consider a special case of the problem on arbitrary group representations, specifcally we consider natural actions of symmetric groups on vector spaces over finite fields. In order to bound the codimension of minimal symmetrically covering subspaces we employ very different methods to the cyclic case.

2.2 Simple upper and lower bounds

We first recall a straightforward lower bound on $h_2(n)$, due to Cameron (unpublished). We give a proof for completeness.

Lemma 2.2.1. *For odd positive integers $n > 3$, we have $h_2(n) \geq 2$.*

Proof. Let $U = \mathrm{Span}(S)$, where

$$S = \{1111111...11, 1010000...00, 0001100...00, 0000110...00, 0000011...00, ..., 0000000...11\}.$$

Since S is a linearly independent set of size $n-2$, we have $\mathrm{codim}(U) = 2$. We claim that U is a cyclically covering subspace of $\mathbb{F}_2^n$. Observe that the last $n-4$ elements of S are a basis for the subspace $\{x \in V_n : x_1 = x_2 = x_3 = 0\}$. First let $x \in \mathbb{F}_2^n$ have even Hamming weight. Since x has an odd number of zeros, it has a (cyclic) interval of consecutive zeros, with odd length. In particular, it contains a (cyclic) interval of the form 000 or 101. By cycling x, we may assume that $x_1x_2x_3 = 000$ or $x_1x_2x_3 = 101$. In the first case, x lies in the span of the last $n-4$ elements of S; in the second, $x + 101000\ldots0$ lies in the span of the last $n-4$ elements of S, so we are done. Now let $x \in \mathbb{F}_2^n$ have odd Hamming weight. Then $x + 11\ldots1$ has even Hamming weight, and $11\ldots1 \in S$, so again we are done. □

It is easy to check that $h_2(3) = 1$, and therefore the assumption $n > 3$ in Lemma 2.2.1 is necessary. Equality holds in Lemma 2.2.1 for $n = 5$.

We next give a rather crude 'product' bound.

Lemma 2.2.2. *If q is a prime power, and $n, m \in \mathbb{N}$, then*

$$h_q(nm) \geq \max\{h_q(n), h_q(m)\}.$$

Remark: We note that Aaronson, Groenland and Johnston [1] have since proven strengthened versions of Lemma 2.2.2, and the related Lemma 2.4.5, showing that

$$h_q(nm) \geq h_q(n) + h_q(m),$$

but we include these results for completeness.

Proof of Lemma 2.2.2. Let q be a prime power. If v is a vector in $\mathbb{F}_q^N$ for some $N \in \mathbb{N}$,

let us write $v(j)$ for the jth component of v (i.e., $v = \sum_{i=1}^{N} v(i)e_i$ with respect to the standard basis $\{e_1, e_2, ..., e_N\}$).

Let $n, m \in \mathbb{N}$. Without loss of generality, we may assume that $h_q(n) \geq h_q(m)$. Let $U \leq \mathbb{F}_q^n$ be a cyclically covering subspace of $\mathbb{F}_q^n$, with $\text{codim}(U) = h_q(n)$. Let $k = h_q(n)$. Let $\{u_1, ..., u_{n-k}\}$ be a basis for U. For each $i \in [n-k]$, let

$$x_i = (\underbrace{0, 0, ..., 0, u_i(1)}_{m}, \underbrace{0, 0, ..., 0, u_i(2)}_{m}, ..., \underbrace{0, 0, ..., 0, u_i(n)}_{m}) \in \mathbb{F}_q^{nm}.$$

Let

$$S = \{x_i : \ i \in [n-k]\} \cup \{e_j : \ j \in [nm], \ m \nmid j\},$$

where e_j is the jth standard basis vector in $\mathbb{F}_q^{nm}$, and let $V = \text{Span}(S)$. Then $|S| = (n-k) + (nm-n) = nm-k$, and S is linearly independent, so $\text{codim}(V) = k$. We claim that V cyclically covers $\mathbb{F}_q^{nm}$. Indeed, if $x \in \mathbb{F}_q^{nm}$, then consider the projection of x onto the subspace spanned by $\{e_j : \ m \mid j\}$, i.e.,

$$\pi(x) := (\underbrace{0, 0, \ldots, 0, x(m)}_{m}, \underbrace{0, 0, \ldots, 0, x(2m)}_{m}, \underbrace{0, 0, \ldots, 0, x(3m)}_{m}, \ldots \underbrace{0, 0, \ldots, 0, x(nm)}_{m}) \in \mathbb{F}_q^{nm},$$

and let

$$\psi(x) := (x(m), x(2m), x(3m), ..., x(nm)) \in \mathbb{F}_q^n$$

be the vector obtained from $\pi(x)$ by deleting the coordinates that are not multiples of m. Since U cyclically covers $\mathbb{F}_q^n$, there exists $r \in [n-1]$ such that $\sigma^r(\psi(x)) \in U$. It follows that $\sigma^{mr}(\pi(x)) \in V$, and therefore $\sigma^{mr}(x) \in V$, since S contains every unit vector e_j such that $m \nmid j$. Hence, V is cyclically covering, as claimed, and therefore $h_q(nm) \geq \text{codim}(V) = k = h_q(n)$, proving the lemma.

□

We now give a straightforward upper bound for $h_q(n)$, for all $n \in \mathbb{N}$.

Lemma 2.2.3. *For q a prime power and $n \in \mathbb{N}$, we have $h_q(n) \leq \lfloor \log_q(n) \rfloor$.*

Proof. Let $U \leq \mathbb{F}_q^n$ be a cyclically covering subspace. The cyclic group

$$\langle \sigma \rangle = \{\mathrm{Id}, \sigma, \sigma^2, \dots \sigma^{n-1}\}$$

acts on $\mathbb{F}_q^n$ by cyclically shifting vectors. The orbits of this group action partition $\mathbb{F}_q^n$, and each orbit contains at most n vectors, so there are at least q^n/n orbits. Since U is cyclically covering, it intersects each orbit, and therefore $|U| \geq q^n/n$. Hence, $\dim(U) = \log_q(|U|) \geq n - \log_q(n)$, so $\mathrm{codim}(U) \leq \log_q(n)$, proving the lemma. □

Our main results show that for each prime power q, the simple upper bound in Lemma 2.2.3 is tight for infinitely many values of n. The proofs of these results occupy most of the next section.

2.3 Our main results

Before getting into the technical details of our main results we give a brief overview of the main ideas behind their proof. In each case the aim is to prove lower bounds on $h_q(n)$ for prime powers q and certain sequences of positive integers n in order to show that the upper bound $h_q(n) \leq \lfloor \log_q(n) \rfloor$ can be tight. In order to prove a lower bound on $h_q(n)$ we find cyclically covering subspaces $U \leq \mathbb{F}_q^n$ and calculate their codimension, since by definition $h_q(n) \geq \mathrm{codim}(U)$.

To find these cyclically covering subspaces we reformulate the problem by noting the equivalence of $\mathbb{F}_q^n$, equipped with the linear map σ, with the polynomial ring $\mathbb{F}_q[X]/\langle X^n - 1 \rangle$ (where the action of σ corresponds to multiplication by X). A cyclically covering subspace $U \leq \mathbb{F}_q[X]/\langle X^n - 1 \rangle$ is then a subspace such that $\bigcup_{i=0}^{n-1} X^i U = \mathbb{F}_q[X]/\langle X^n - 1 \rangle$. This reformulation helps us identify natural subspaces that are invariant under multiplication by X by factorising $X^n - 1$. It is useful to find invariant subspaces since a cyclically covering subspace of an invariant subspace can be easily extended to a cycli-

cally covering subspace of the whole space. Hence if we are able to identify a "particularly nice" invariant subspace, in a sense we will define presently, we might hope that finding a cyclically covering subspace of this nice subspace will provide a good bound for the whole space and indeed this is what happens.

By "particularly nice" we mean that the reasoning behind our log bound is the whole picture, i.e., the orbit structure under cyclic action is sufficiently simple that to construct a cyclically covering subspace it is sufficient to quite crudely pick elements of the orbits. For example, in our first main theorem there are only two orbits in the invariant subspace: one singleton containing 0 and the other containing all non-zero elements. It is then clear that picking the subspace generated by any single non-zero element will be cyclically covering and as small as possible. In order to identify "particularly nice" invariant subspaces we will use some standard results from the theory of finite fields to harness the algebraic properties of carefuly chosen roots of $X^n - 1$ in the algebraic closure of $\mathbb{F}_q^n$.

Our first main result is as follows.

Theorem 2.3.1. *If q is a prime power and $d \in \mathbb{N}$, then*

$$h_q(q^d - 1) = d - 1 = \left\lfloor \log_q(q^d - 1) \right\rfloor .$$

Observe that the upper bound $h_q(q^d - 1) \leq d - 1$ is immediate from Lemma 2.2.3. Our proof of the lower bound $h_q(q^d - 1) \geq d - 1$ requires some standard facts from the Galois theory of finite fields, which we now briefly recall; the reader is referred to [58] for more background.

For a prime power q, we write $\overline{\mathbb{F}}_q$ for the *algebraic closure* of $\mathbb{F}_q$, i.e., the algebraic field extension of $\mathbb{F}_q$ that is algebraically closed. We write $F_q : \overline{\mathbb{F}}_q \to \overline{\mathbb{F}}_q;\ x \mapsto x^q$ for the *Frobenius automorphism* of $\overline{\mathbb{F}}_q$. Note that if $a \in \mathbb{F}_q \setminus \{0\}$ then $a^{q-1} = 1$: indeed as a is

non-zero we see multiplication by a is a bijection on $\mathbb{F}_q$ fixing 0, and so

$$a^{q-1} \cdot \prod_{b \in \mathbb{F}_q \setminus \{0\}} b = \prod_{b \in \mathbb{F}_q \setminus \{0\}} (ab) = \prod_{b \in \mathbb{F}_q \setminus \{0\}} b.$$

Noting $\prod_{b \in \mathbb{F}_q \setminus \{0\}} b$ is non-zero and so invertible proves the result.

For $\omega \in \overline{\mathbb{F}}_q$, we call $\omega, F_q(\omega), F_q^2(\omega), F_q^3(\omega), \ldots$ the *Galois conjugates* of ω; if $\omega \in \overline{\mathbb{F}}_q$ is a root of some polynomial $f \in \mathbb{F}_q[X]$, then all the Galois conjugates of ω are also roots of f. Indeed suppose that

$$f(X) = \sum_{i=0}^{k} a_i X^i,$$

where $a_i \in \mathbb{F}_q$ for $i = 1, ..., k$. If q is a power of prime p then as $\mathbb{F}_q$ is a field extension (and hence a vector space) over $\mathbb{F}_p$ we have $p \cdot a = 0$ for all $a \in \mathbb{F}_q$. If $k = 0$ it is trivially true that $f(X)^p = {a_0}^p$. If $k \geq 1$, set $g(X) = \sum_{i=0}^{k-1} a_i X^i$. Using the well known fact that $p \mid \binom{p}{r}$ for $r = 2, 3, ..., p-1$, we have

$$f(X)^p = \sum_{r=0}^{p} \binom{p}{r} g(X)^r {a_k}^{p-r} X^{k(p-r)} = g(X)^p + {a_k}^p X^{kp} = \sum_{i=0}^{k} {a_i}^p X^{ip},$$

where the final equality is by induction on k. It follows that

$$f(X)^q = \sum_{i=0}^{k} {a_i}^q X^{iq} = \sum_{i=0}^{k} a_i X^{iq} = f(X^q),$$

where in the second equality we are using that $a^q = a$ for $a \in \mathbb{F}_q$, as proven above. Hence, if $\omega \in \overline{\mathbb{F}_q}$ such that $f(\omega) = 0$, then $f(\omega^q) = f(\omega)^q = 0$.

Since any element $\omega \in \overline{\mathbb{F}}_q$ is the root of some polynomial in $\mathbb{F}_q[X]$, and all of the Galois conjugates of ω are roots of this polynomial, any $\omega \in \overline{\mathbb{F}}_q$ has only finitely many Galois conjugates. If $\omega \in \overline{\mathbb{F}}_q$, the *minimal polynomial* of ω over $\mathbb{F}_q$ is the unique non-zero, monic polynomial in $\mathbb{F}_q[X]$ of minimal degree, that has ω as a root.

We will make repeated use of the following well-known fact (see for example Theorem 3.33 in [58]).

Proposition 2.3.1. *Let q be a prime power and let $\omega \in \overline{\mathbb{F}}_q$. Let $\omega, \omega^q, \ldots, \omega^{q^{t-1}}$ be the distinct Galois conjugates of ω. Then*

$$f(X) = \prod_{i=1}^{t} (X - \omega^{q^{t-1}})$$

is the minimal polynomial of ω.

We are now ready to prove Theorem 2.3.1.

Proof of Theorem 2.3.1. Let q be a prime power, let $d \in \mathbb{N}$ and let $n = q^d - 1$. We identify $\mathbb{F}_q^n$ and $\mathbb{F}_q[X]/\langle X^n - 1\rangle$ via the linear isomorphism taking $v \in \mathbb{F}_q^n$ to the polynomial $\sum_{i=1}^{n} v(i)X^{i-1}$. The action of σ on $\mathbb{F}_q^n$ then corresponds to multiplication by X in $\mathbb{F}_q[X]/\langle X^n - 1\rangle$.

Since $X^n - 1$ and $nX^{n-1} = \frac{d}{dX}(X^n - 1)$ are coprime, it follows that $X^n - 1$ has no repeated roots in $\overline{\mathbb{F}}_q$. Let

$$\prod_{i=1}^{N} f_i(X) = X^n - 1 \tag{2.2}$$

be a factorisation of $X^n - 1$ into monic irreducible polynomials $f_i(X) \in \mathbb{F}_q[X]$. Since $X^n - 1$ has no repeated roots in $\overline{\mathbb{F}}_q$, the $f_i(X)$ are distinct. Moreover, each pair $f_i(X), f_j(X)$ is coprime, since if $p(X) \neq 1$ is a monic common factor of $f_i(X)$ and $f_j(X)$ then by irreducibility, we have $p(X) = f_i(X) = f_j(X)$, and therefore $i = j$.

Define a linear map

$$\theta : \frac{\mathbb{F}_q[X]}{\langle X^n - 1\rangle} \to \bigoplus_{i=1}^{N} \frac{\mathbb{F}_q[X]}{\langle f_i(X)\rangle}; \quad \theta(p(X)) = (p(X) \bmod f_i(X))_{i=1}^{N},$$

i.e., θ is the direct sum of the natural quotient maps corresponding to the ideals generated by each f_i. Since the $f_i(X)$ are pairwise coprime, it follows from the Chinese Remainder

Theorem for rings that θ is a linear isomorphism. For each $i \in [N]$, define

$$V_i = \left\{ p(X) \in \frac{\mathbb{F}_q[X]}{\langle X^n - 1 \rangle} \ : \ \prod_{j \neq i} f_j(X) \text{ divides } p(X) \right\}.$$

Since for each $i \in [N]$, we have

$$V_i = \theta^{-1} \left(\{0\} \times \ldots \times \{0\} \times \frac{\mathbb{F}_q[X]}{\langle f_i(X) \rangle} \times \{0\} \times \ldots \times \{0\} \right),$$

(where the zeros are in each place except for the ith) and θ is a linear isomorphism, we have the direct sum decomposition

$$\frac{\mathbb{F}_q[X]}{\langle X^n - 1 \rangle} = \bigoplus_{i=1}^{N} V_i, \tag{2.3}$$

and V_i may be viewed as the copy of $\mathbb{F}_q[X]/\langle f_i(X) \rangle$ in $\mathbb{F}_q[X]/\langle X^n - 1 \rangle$, for each $i \in [N]$. Moreover, each V_i is closed under multiplication by X (i.e., under the cyclic action of σ the V_i are *invariant* subspaces).

Since $\text{char}(\mathbb{F}_q) \nmid n$, there exists a primitive nth root of unity in $\overline{\mathbb{F}}_q$. Let $\omega \in \overline{\mathbb{F}}_q$ be one such. Since $n = q^d - 1$, q has multiplicative order d modulo n, and so the iterates of ω under the Frobenius automorphism are precisely $\omega, \omega^q, \omega^{q^2}, \ldots, \omega^{q^{d-1}}$ (and these are distinct). Hence, by Proposition 2.3.1, the minimal polynomial of ω over $\mathbb{F}_q$ is

$$f(X) = (X - \omega)(X - \omega^q)(X - \omega^{q^2})...(X - \omega^{q^{d-1}}) \in \mathbb{F}_q[X]. \tag{2.4}$$

As $f(X)$ is a monic irreducible factor of $X^n - 1$, we may take $f_1(X) = f(X)$ in the factorisation (2.2).

Let $u(X) \in V_1 \leq \mathbb{F}_q[X]/\langle X^n - 1 \rangle$ such that $u(X) \equiv 1 \bmod f(X)$. We claim that the cyclic orbit of $u(X)$ (i.e., its orbit under repeated multiplication by X) is equal to $V_1 \setminus \{0\}$.

To prove this, we first observe that $X^m u(X) \neq u(X)$ for all $1 \leq m \leq n-1$. Indeed,

suppose for a contradiction there exists $m \in [n-1]$ such that multiplication by X^m fixes $u(X)$ in $\mathbb{F}_q[X]/\langle X^n - 1\rangle$. Then $X^m u(X) \equiv u(X) \bmod (X^n - 1)$, and therefore $X^n - 1$ divides $(X^m - 1)u(X)$. It follows that ω is a root of $(X^m - 1)u(X)$. Since ω is a primitive nth root of unity, we have $\omega^m - 1 \neq 0$, and therefore $u(\omega) = 0$. Hence, as $f(X)$ is the minimal polynomial of ω, $f(X)$ divides $u(X)$. This contradicts our assumption that $u(X) \equiv 1 \bmod f(X)$.

It follows that $u(X), Xu(X), X^2u(X), ..., X^{n-1}u(X)$ are n distinct elements of $V_1 \setminus \{0\}$. Since $\dim(V_1) = \deg(f(X)) = d$, we have $|V_1 \setminus \{0\}| = q^d - 1 = n$. It follows that the cyclic orbit of $u(X)$ is precisely $V_1 \setminus \{0\}$, as claimed.

Let $U = \mathrm{Span}\{u(X)\} \leq V_1$. Clearly, by the preceding claim, U cyclically covers V_1. Note that the codimension of U as a subspace of V_1 is $\dim(V_1) - 1 = d - 1$.

Finally, we set

$$U' = U \oplus \left(\bigoplus_{i=2}^{N} V_i \right) \leq \frac{\mathbb{F}_q[X]}{\langle X^n - 1\rangle},$$

and we claim that U' cyclically covers $\mathbb{F}_q[X]/\langle X^n - 1\rangle$. Indeed, given $v(X) \in \mathbb{F}_q[X]/\langle X^n - 1\rangle$, there exist unique $v_i(X) \in V_i$ (for $i = 1, 2, \ldots, N$) such that $v(X) = \sum_{i=1}^{N} v_i(X)$. Since U cyclically covers V_1, there exists $m \in \{0, 1, 2, ..., n-1\}$ such that $X^m v_1(X) \in U$. Since $X^m v_i(X) \in V_i$ for all $i \in [N]$, it follows that $X^m v(X) = \sum_{i=1}^{N} X^m v_i(X) \in U'$. Hence, U' cyclically covers $\mathbb{F}_q[X]/\langle X^n - 1\rangle$. Since the codimension of U in V_1 is equal to $d - 1$, the codimension of U' in $\mathbb{F}_q[X]/\langle X^n - 1\rangle$ is also equal to $d - 1$.

It follows that $h_q(n) \geq d - 1$. In combination with the upper bound in Lemma 2.2.3, this completes the proof of the theorem.

□

With only a little extra work, the argument in the proof of Theorem 2.3.1 can be extended to obtain the following more general lower bound. Here we will also make use of Bézout's lemma for polynomials which we recall for completeness.

Lemma 2.3.1 (Bézout's lemma for polynomials)**.** *Let* $\mathbb{F}$ *be a field and* $p(X), q(X) \in \mathbb{F}[X]$ *be polynomials with greatest common divisor* $d(X)$ *(i.e., for all* $f(X)$ *such that* $f(X) \mid p(X)$ *and* $f(X) \mid q(X)$ *we have* $f(X) \mid d(X)$*). Then there exist polynomials* $s(X), t(X) \in \mathbb{F}[X]$ *such that*

$$s(X)p(X) + t(X)q(X) = d(X).$$

Proof. Without loss of generality the degree of $p(X)$ is at least the degree of $q(X)$. By polynomial division, we find $u_0(X), v_0(X) \in \mathbb{F}[X]$ such that

$$p(X) = v_0(X)q(X) + u_0(X),$$

and the degree of $u_0(X)$ is strictly smaller than the degree of $q(X)$. Note also that $d(X) \mid u_0(X)$. If $u_0(X) = 0$, then $d(X) = q(X)$ and we are done, so without loss of generality $u_0(X) \neq 0$. Again, by polynomial division, we find $u_1(X), v_1(X) \in \mathbb{F}[X]$ such that

$$q(X) = v_1(X)u_0(X) + u_1(X),$$

and the degree of $u_1(X)$ is strictly smaller than the degree of $u_0(X)$. Note also that $d(X) \mid u_1(X)$. Proceeding inductively, we find $u_i(X), v_i(X) \in \mathbb{F}[X]$ such that

$$u_{i-2}(X) = v_i(X)u_{i-1}(X) + u_i(X)$$

and the degree of $u_i(X)$ is strictly smaller than that of $u_{i-1}(X)$. Since $d(X) \mid u_{i-2}(X)$ and $d(X) \mid u_{i-1}(X)$ we see that $d(X) \mid u_i(X)$. We halt this process when $u_{i+1}(X) = 0$.

Then $v_{i+1}(X)u_i(X) = u_{i-1}(X)$, so $u_i(X) \mid u_{i-1}(X)$, and $u_i(X) + v_i(X)u_{i-1}(X) = u_{i-2}(X)$, so $u_i(X) \mid u_{i-2}(X)$. Suppose that $u_i(X) \mid u_{i-r+2}$ and $u_i(X) \mid u_{i-r+1}(X)$, then since

$$u_{i-r}(X) = v_{i-t+2}(X)u_{i-r+1}(X) + u_{i-r+2}(X),$$

we see that $u_i(X) \mid u_{i-r}(X)$. Hence, by induction on r we have that $u_i(X) \mid p(X)$ and

$u_i(X) \mid q(X)$, and so $u_i(X) \mid d(X)$.

Now, since $d(X) \mid u_i(X)$ and $u_i(X) \mid d(X)$, we deduce that (up to multiplication by an invertible element in $\mathbb{F}$) $d(X) = u_i(X)$. We now read the equalities above in reverse to find $s(X)$ and $t(X)$. Indeed

$$u_i(X) = u_{i-2}(X) - v_i(X)u_{i-1}(X),$$

so set $s_1(X) = 1$ and $t_1(X) = -v_i(X)$ and supposing that

$$u_i(X) = s_r(X)u_{i-r-1}(X) + t_r(X)u_{i-r}(X),$$

then since

$$u_{i-r-2}(X) = v_{i-r}(X)u_{i-r-1}(X) + u_{i-r}(X),$$

we may set $s_{r+1}(X) = t_r(X)$ and $t_{r+1}(X) = s_r(X) - v_{i-r}(X)t_r(X)$, and find $s(X), t(X)$ by induction on r. □

Theorem 2.3.2. *Let q be a prime power and let $k, d \in \mathbb{N}$. Let $M = (q-1)(\sum_{r=0}^{d} q^{kr}) = (q-1)(q^{kd+k}-1)/(q^k-1)$, and suppose that M has a divisor $c \in \mathbb{N}$ such that $c < (q-1)\frac{q^k - q^{-kd}}{q^k-1}$. Then*

$$h_q(M/c) \geq kd + k - c(q^k-1)/(q-1).$$

Proof. Let q be a prime power, let $k, d \in \mathbb{N}$ and let $M = (q-1)(\sum_{r=0}^{d} q^{kr})$. Let $c \in \mathbb{N}$ be a divisor of M satisfying $c < (q-1)\frac{q^k - q^{-kd}}{q^k-1}$. Set $n := M/c$. As in the proof of Theorem 2.3.1, we identify $\mathbb{F}_q^n$ with $\mathbb{F}_q[X]/\langle X^n - 1\rangle$, and we decompose the latter into invariant subspaces,

$$\mathbb{F}_q[X]/\langle X^n - 1\rangle = \bigoplus_{i=1}^{N} V_i,$$

by taking a factorisation

$$X^n - 1 = \prod_{i=1}^{N} f_i(X)$$

of $X^n - 1$ into a product of irreducible monic factors, and taking V_i to be the preimage of $\mathbb{F}_q[X]/\langle f_i(X)\rangle$ under the direct sum of the natural quotient maps, $p(X) \mapsto p(X) \bmod f_i(X)$.

Note that the V_i are irreducible subspaces. Indeed, suppose that $\{0\} \neq W \leq V_i$ and that W is invariant under multiplication by X. Let $p(X) \in \mathbb{F}_q[X]$ such that the image of p (under the natural quotient map $\mathbb{F}_q[X] \to \mathbb{F}_q[X]/\langle X^n - 1\rangle$) lies in $W \setminus \{0\}$. Then $f_i(X)$ does not divide $p(X)$, and $f_i(X)$ is irreducible, so $p(X)$ and $f_i(X)$ are coprime. Hence, by Bézout's lemma, there exist $s(X), t(X) \in \mathbb{F}_q[X]$ such that $s(X)p(X) + t(X)f_i(X) = 1$. Let $q(X) := s(X)p(X)$; we have $q(X) \equiv 1 \bmod f_i(X)$. The invariance of W under multiplication by X implies that $q(X) \in W$, and moreover that

$$\{q(X), Xq(X), X^2q(X), \ldots\} \subseteq W$$

But the set on the left-hand side contains a basis for V_i, since for each $0 \leq r \leq n-1$ we have $X^r q(X) \equiv X^r \bmod f_i(X)$. It follows that that $W = V_i$, proving the irreducibility of V_i.

We now continue to follow the proof of Theorem 2.3.1. Let $\omega \in \overline{\mathbb{F}}_q$ be a primitive nth root of unity. We claim that the order of q modulo n is $k(d+1)$. Indeed, let L be the order of q modulo n. Since $nc(\sum_{t=0}^{k-1} q^t) = q^{k(d+1)} - 1$, we have $q^{k(d+1)} \equiv 1 \bmod n$, and therefore L divides $k(d+1)$. Since $q^{kd} < n \Leftrightarrow c < (q-1)\frac{q^k - q^{-kd}}{q^k - 1}$, we have $L > kd$. Since $kd \geq \frac{1}{2}k(d+1)$, no non-trivial factor of $k(d+1)$ is greater than kd, and therefore $L = k(d+1)$, as claimed. By Proposition 2.3.1, the minimal polynomial of ω over $\mathbb{F}_q$ is

$$f(X) = (X - \omega)(X - \omega^q)(X - \omega^{q^2})...(X - \omega^{q^{k(d+1)-1}}) \in \mathbb{F}_q[X],$$

which has degree $k(d+1)$. We may assume without loss of generality that $f_1(X) = f(X)$,

and consider V_1. As in the proof of Theorem 2.3.1, let $u(X) \in V_1 \leq \mathbb{F}_q[X]/\langle X^n - 1\rangle$ such that $u(X) \equiv 1 \bmod f_1(X)$, and recall from the proof of Theorem 2.3.1 that $u(X)$ has orbit (under repeated multiplication by X) of size exactly n. More generally, let $0 \neq v(X) \in V_1$; we claim that the orbit

$$\{v(X), Xv(X), X^2v(X), ..., X^{n-1}v(X)\} \subseteq V_1$$

also has size exactly n. Indeed, since V_1 is irreducible, and

$$0 \neq \text{Span}(\{v(X), Xv(X), X^2v(X), ..., X^{n-1}v(X)\}) \leq V_1$$

is a subspace that is invariant under multiplication by X, we see that

$$\{v(X), Xv(X), X^2v(X), ..., X^{n-1}v(X)\}$$

spans V_1. Suppose for a contradiction there exists $1 \leq a \leq n-1$ such that $X^a v(X) = v(X)$ (note that this is an equality in $\mathbb{F}_q[X]/\langle X^n - 1\rangle$). We may then express $u(X)$ as a linear combination $u(X) = \sum_{i=0}^{a-1} \lambda_i X^i v(X)$, for some $\lambda_i \in \mathbb{F}_q$. But then

$$X^a u(X) = \sum_{i=0}^{a-1} \lambda_i X^i X^a v(X) = \sum_{i=0}^{a-1} \lambda_i X^i v(X) = u(X)$$

contradicting the fact that the orbit of $u(X)$ has size exactly n. It follows that $v(X)$ has orbit of size exactly n, as claimed.

We may conclude all the orbits (under repeated multiplication by X) in $V_1 \setminus \{0\}$ have size n. There are $s := (|V_1| - 1)/n = (q^{k(d+1)} - 1)/n = c\sum_{t=0}^{k-1} q^t$ such orbits; let $\{u_1, u_2, ..., u_s\}$ be a set of representatives of these orbits. Then $U = \text{Span}(\{u_1, u_2, ..., u_s\}) \leq V_1$ cyclically covers V_1, and has codimension (in V_1) at least $k(d+1) - s$.

Finally, we set

$$U' = U \oplus \left(\bigoplus_{i \neq 1} V_i \right) \leq \frac{\mathbb{F}_q[X]}{\langle X^n - 1\rangle};$$

note that U' cyclically covers $\mathbb{F}_q[X]/\langle X^n-1\rangle$ and has codimension (in $\mathbb{F}_q[X]/\langle X^n-1\rangle$) at least $k(d+1)-s = kd+k-c\sum_{t=0}^{k-1} q^t = kd+k-c(q^k-1)/(q-1)$. It follows that $h_q(n) \geq kd+k-c(q^k-1)/(q-1)$, as required.

□

Applying the above theorem with $c=1$, fixed q,k and $d\to\infty$, and appealing to Lemma 2.2.3, we see that

$$h_q\left((q-1)\sum_{r=0}^{d} q^{kr}\right) = (1+o(1))kd$$

where the $o(1)$ term tends to zero as d tends to infinity. Theorem 2.3.1 is recovered from Theorem 2.3.2 by setting $k=1$ and $c=1$.

We now demonstrate how a slight variation on the ideas in the proofs of Theorem 2.3.1 and Theorem 2.3.2 can determine $h_q(n)$ for other infinite sequences of integers n (for each fixed prime power q).

Theorem 2.3.3. *Let q be a prime power, and let $k,d\in\mathbb{N}$ such that $gcd(d+1,q^k-1)=1$. Set $n=\sum_{r=0}^{d} q^{kr} = \frac{q^{k(d+1)}-1}{q^k-1}$. Then*

$$h_q(n) = kd.$$

Proof. The upper bound $h_q(n)\leq kd$ follows immediately from Lemma 2.2.3, so we need only prove the lower bound. We first note that

$$n=\sum_{r=0}^{d} q^{kr} \equiv d+1 \bmod q^k-1,$$

since $q^{kr}\equiv 1 \bmod q^k-1$ for each $r\in\mathbb{N}\cup\{0\}$, so

$$\gcd(n,q^k-1) = \gcd(d+1,q^k-1) = 1. \tag{2.5}$$

As in the proofs of Theorems 2.3.1 and 2.3.2, we identify $\mathbb{F}_q^n$ with $\mathbb{F}_q[X]/\langle X^n-1\rangle$, and we decompose the latter into invariant subspaces,

$$\mathbb{F}_q[X]/\langle X^n-1\rangle = \bigoplus_{i=1}^{N} V_i,$$

by taking a factorisation

$$X^n - 1 = \prod_{i=1}^{N} f_i(X)$$

of X^n-1 into a product of irreducible monic factors, and taking V_i to be the preimage of $\mathbb{F}_q[X]/\langle f_i(X)\rangle$ under the direct sum of the natural quotient maps.

Let $\omega \in \overline{\mathbb{F}}_q$ be a primitive nth root of unity. As in the proof of Theorem 2.3.2, we claim that q has multiplicative order $k(d+1)$ modulo n. Indeed, let L be the order of q modulo n. Since $q^{k(d+1)}-1 = n(q^k-1)$, we have $q^{k(d+1)} \equiv 1 \bmod n$, and therefore L divides $k(d+1)$. Since $q^{kd} < n$, we must have $L > kd$. Since $kd \geq \frac{1}{2}k(d+1)$, no non-trivial factor of $k(d+1)$ is greater than kd, and therefore $L = k(d+1)$, as claimed. By Proposition 2.3.1, the minimal polynomial of ω over $\mathbb{F}_q$ is

$$f(X) = (X-\omega)(X-\omega^q)(X-\omega^{q^2})...(X-\omega^{q^{k(d+1)-1}}) \in \mathbb{F}_q[X],$$

which has degree $k(d+1)$. We may assume without loss of generality that $f_1(X) = f(X)$, and consider V_1. Since $V_1 \cong \mathbb{F}_q[X]/\langle f(X)\rangle$ and $f(X)$ is an irreducible polynomial of degree $k(d+1)$, V_1 is in fact a field extension of $\mathbb{F}_q$ (of degree $k(d+1)$), and as such can be identified with the finite field $\mathbb{F}_{q^{k(d+1)}}$. Hence, V_1 can also be viewed as a $(d+1)$-dimensional vector space over (a field isomorphic to) $\mathbb{F}_{q^k}$.

Let U be a 1-dimensional $\mathbb{F}_{q^k}$-subspace of V_1. We now make two claims regarding U. Firstly, we claim that no power of the shift map can map U to itself. Indeed, suppose for a contradiction that there exists $a \in [n-1]$ such that $X^aU = U$. Then for any $u \in U \setminus \{0\}$, we have

$$\{u, X^a u, X^{2a} u, \ldots\} \subseteq U.$$

We note, as in the proof of Theorem 2.3.2, that for every $v \in V_1 \setminus \{0\}$, the orbit of v under repeated multiplication by X has size n. Hence, for $j \in \mathbb{N}$, $X^j v = v$ if and only if $n \mid j$. It follows that for any $u \in U \setminus \{0\}$, we have $X^{aj}u = u$ if and only if $n \mid aj$, i.e., if and only if $n/\gcd(a,n) \mid j$. Therefore, the above orbit of u under repeated multiplication by X^a has size exactly $n/\gcd(a,n) := M$. The family of all such orbits (of non-zero elements of U, under repeated multiplication by X^a) partitions $U \setminus \{0\}$ into sets of equal size M, and therefore M is a proper divisor of n that also divides $|U| - 1 = q^k - 1$. But this contradicts (2.5).

Secondly, we claim that $X^b U \cap X^c U = \{0\}$ for any $0 \leq b < c \leq n-1$. Indeed, suppose for a contradiction that there exist $0 \leq b < c \leq n-1$ such that $X^b U \cap X^c U \neq \{0\}$. Then, multiplying by X^{n-b}, we have $X^n U \cap X^{n+c-b} U \neq \{0\}$ and therefore $U \cap X^a U \neq \emptyset$, where $a := c - b \in [n-1]$. However, U and $X^a U$ are distinct 1-dimensional $\mathbb{F}_{q^k}$-subspaces of V_1, so have intersection $\{0\}$, a contradiction.

It follows that $U, XU, \ldots, X^{n-1}U$ are q^k-element subsets of V_1 whose pairwise intersections are all equal to $\{0\}$; since $q^{k(d+1)} - 1 = n(q^k - 1)$, we must have $V_1 = \cup_{a=0}^{n-1} X^a U$, so U (as an $\mathbb{F}_q$-subspace of V_1) is a cyclic cover of V_1 with codimension kd in V_1.

Set $U' = U \oplus (\oplus_{i \neq 1} V_i)$; then U' is a cyclic cover of $\mathbb{F}_q[X]/\langle X^n - 1 \rangle$ with codimension kd, so $h_q(n) \geq kd$, as required. □

We note that for a fixed prime power q and a fixed integer k, a positive fraction of positive integers d have the property that $\gcd(d+1, q^k - 1) = 1$ (and so satisfy the hypothesis of Theorem 2.3.3).

Implications for the lower bound in Isbell's conjecture and the generalised Isbell conjecture

The $q = 2$ cases of Theorems 2.3.1 and 2.3.3, together with the $p = 2$ case of Proposition 2.1.1, imply the following.

Corollary 2.3.1. *For any $d \in \mathbb{N}$, we have $h_2(2^d-1) = d-1$, and therefore $m(2^d-1) \geq d$. Moreover, for any $d, k \in \mathbb{N}$ with $gcd(d+1, 2^k-1) = 1$, we have $h_2((2^{k(d+1)}-1)/(2^k-1)) = kd$, and therefore $m((2^{k(d+1)}-1)/(2^k-1)) \geq kd+1$.*

We remark that a construction due to Spiga [69] (building on the work of Suzuki [70] in which he introduced and analysed the Suzuki groups), gives

$$m((2^{kr}-1)/(2^k-1)) \geq k(r-1)^2+1$$

for all primes $r > 2$ and integers $k \in \mathbb{N}$ coprime to 2^r-1. In particular, $m(2^r-1) \geq (r-1)^2+1$ for all primes $r > 2$, giving a lower bound on $m(b)$ that is quadratic in $\log b$ for infinitely many b, whereas our lower bound in Corollary 2.3.1 is only linear in $\log b$. Spiga's construction involves replacing the Abelian group V (in the penultimate paragraph of the Introduction) with a non-Abelian group N (for example, a certain Sylow 2-subgroup of a Suzuki group), and finding a subgroup of N of large index, whose images under an appropriate cyclic automorphism cover N. However, our construction is simpler and provides good lower bounds on the function m for other natural infinite sequences of odd integers, where Spiga's method does not apply.

Similarly, the $q = p$ cases of Theorems 2.3.1 and 2.3.3, together with the general case of Proposition 2.1.1, imply the following.

Corollary 2.3.2. *For any $d \in \mathbb{N}$, we have $h_p(p^d-1) = d-1$, and therefore $m_p(p^d-1) \geq d$. Moreover, for any $d, k \in \mathbb{N}$ with $gcd(d+1, p^k-1) = 1$, we have $h_p((p^{k(d+1)}-1)/(p^k-1)) = kd$, and therefore $m_p((p^{k(d+1)}-1)/(p^k-1)) \geq kd+1$.*

In this case, Spiga's construction in [69] yields

$$m_p((p^{kr}-1)/(p^k-1)) \geq k(r-1)^2+1$$

for all primes $r \neq p$ and $k \in \mathbb{N}$ such that r and p^k-1 are coprime. In particular, $m_p((p^r-1)/(p-1)) \geq (r-1)^2+1$ for all primes $r \neq p$ such that r does not divide $p-1$.

Again, for a fixed prime p, this gives a lower bound on $m_p(b)$ that is quadratic in $\log b$ for infinitely many b, whereas our lower bound in Corollary 2.3.2 is only linear in $\log b$; but again, our construction is simpler and works in cases where Spiga's method does not apply.

2.4 General representations of groups

In this section, we generalise our discussion to arbitrary group representations. The proofs of the bounds in this section are straightforward, given some basic facts from the representation theory of finite groups; nevertheless, each bound is tight in some non-trivial cases.

Let G be a group, let $\mathbb{F}$ be a field and let V be a vector space over $\mathbb{F}$. We write $\mathrm{GL}(V)$ for the general linear group of V. Let $\rho : G \to \mathrm{GL}(V)$ be a group homomorphism, i.e., (ρ, V) is a representation of G. Let us say that a subspace $U \leq V$ is (G, ρ)-*covering* if

$$\bigcup_{g \in G} \rho(g)(U) = V$$

where $\rho(g)(U) := \{\rho(g)(u) : \ u \in U\}$. Let us define $h_{G,\rho}(V)$ to be the maximum possible codimension of a (G, ρ)-covering subspace of V. Note that $h_q(n) = h_{C_n,\rho_\sigma}(\mathbb{F}_q^n)$ for any prime power q and any $n \in \mathbb{N}$, where $(\rho_\sigma, \mathbb{F}_q^n)$ is the representation of the *cyclic group*

$$C_n := \{e, g, g^2, ..., g^{n-1} : g^i = g^j \text{ if and only if } i \equiv j \mod n\}$$

that maps the generator g of C_n to σ.

Let us briefly outline the representation-theoretic terminology and notation we will use. As usual, from now on we will sometimes write $g(u)$ in place of $\rho(g)(u)$, when the representation ρ is understood. Recall that if (ρ, V) is a fixed representation of G, a subspace $W \leq V$ is said to be G-*invariant* if $\rho(g)(w) \in W$ for any $w \in W$ and any $g \in G$; in this case, (ρ, W) is said to be a *subrepresentation* of (ρ, V). Abusing

terminology slightly, when ρ is understood, we will sometimes omit it from our notation, and describe W as a subrepresentation of V.

We also recall Maschke's theorem, an important result from representation theory, and provide a proof for completeness.

Theorem 2.4.1 (Maschke's Theorem)**.** *Let G be a finite group, let $\mathbb{F}$ be a field such that char$(\mathbb{F}) \nmid |G|$, let V be a finite dimensional vector space over $\mathbb{F}$, and (ρ, V) be a representation of G. If $W \leq V$ is a G-invariant subspace (i.e., for all $w \in W$ and $g \in G$ we have $\rho(g)(w) \in W$), then there exists $W' \leq V$ another G-invariant subspace such that $V = W \oplus W'$.*

Proof. In the following proof we suppress ρ in our notation, writing $g(v)$ instead of $\rho(g)(v)$ for $g \in G$ and $v \in V$. Suppose that $W \leq V$ is a G-invariant subspace, then let $\{w_1, ..., w_k\} \subseteq W$ be a basis for W. Extend this basis to $\{w_1, ..., w_m\} \subseteq V$, a basis for V. Define $\pi : V \to W$ by $\pi\left(\sum_{i=1}^m \lambda_i w_i\right) = \sum_{i=1}^k \lambda_i w_i$, and define

$$\bar{\pi}(w) = \frac{1}{|G|} \sum_{g \in G} g^{-1}(\pi(g(w))).$$

Let $W' = \ker(\bar{\pi})$. The image of $\bar{\pi}$ is W: indeed if $w \in W \leq V$ then

$$\bar{\pi}(w) = \frac{1}{|G|} \sum_{g \in G} g^{-1}(\pi(g(w))) = \frac{1}{|G|} \sum_{g \in G} g^{-1}(g(w)) = w.$$

where for the second equality we are using that W is G-invariant and π fixes W.

Hence $V = W \oplus W'$, and furthermore W' is G-invariant. Indeed if $v \in W'$ and $g \in G$, then $\bar{\pi}(g(v)) = g(\bar{\pi}(v)) = g(0) = 0$, hence $g(v) \in W'$. □

If G is a finite group, q is a prime power, V is a finite dimensional vector space over $\mathbb{F}_q$ and (ρ, V) is a representation of G, it is easy to obtain the upper bound

$$h_{G,\rho}(V) \leq \left\lfloor \log_q |G| \right\rfloor, \tag{2.6}$$

just as in the proof of Lemma 2.2.3, since the group G acts on V, partitioning V into orbits, each of size at most $|G|$.

Turning to general lower bounds, the following is easy to obtain.

Lemma 2.4.1. *Let $\mathbb{F}$ be a field, let G be a finite group such that $char(\mathbb{F}) \nmid |G|$, let V be a finite dimensional vector space over $\mathbb{F}$, and let (ρ, V) be a representation of G. If W is a subrepresentation of V, then*

$$h_{G,\rho}(W) \leq h_{G,\rho}(V).$$

Proof. As in the standard proof of Maschke's theorem, we can find a G-invariant subspace W' of V such that $V = W \oplus W'$. (Let $\{w_1, ..., w_k\}$ be a basis for W, and extend it to a basis $\{w_1, ..., w_m\}$ for V. Define $\pi : V \to W$ by $\pi(\sum_{i=1}^m \lambda_i w_i) = \sum_{i=1}^k \lambda_i w_i$, and define

$$\bar{\pi}(w) = \frac{1}{|G|} \sum_{g \in G} g^{-1}(\pi(g(w))).$$

Let $W' = \ker(\bar{\pi})$; then $V = W \oplus W'$, and W' is G-invariant, since if $v \in W'$ and $g \in G$ then $\bar{\pi}(g(v)) = g(\bar{\pi}(v)) = 0$, so $g(v) \in W'$.

Now let $Z \leq W$ be a (G, ρ)-covering subspace of W with codimension $h_{G,\rho}(W)$, and let $U = Z \oplus W' \leq V$. It is easy to see that U is a (G, ρ)-covering subspace of V; clearly, its codimension in V is the same as that of Z in W. Hence, $h_{G,\rho}(V) \geq \text{codim}(U) = \text{codim}(Z) = h_{G,\rho}(W)$, proving the lemma. □

With a little extra work we can achieve the following lower bound:

Lemma 2.4.2. *Let $\mathbb{F}$ be a field, let G be a finite group, let V be a vector space over $\mathbb{F}$, and let (ρ, V) be a representation of G. Suppose $V = \bigoplus_{i=1}^k W_i$ is a decomposition into invariant subspaces. Then for $i = 1, ..., k$, let $\rho_i : G \to GL(W_i)$ be the restriction of ρ to W_i, let $K_i = ker(\rho_i) \leq G$, and let $G_1 = G$ and $G_i = \bigcap_{j=1}^{i-1} K_j$ for $i = 2, ..., k$.*

Then:

$$\sum_{i=1}^{k} h_{G_i,\rho_i}(W_i) \leq h_{G,\rho}(V).$$

Proof. For each $i = 1, ..., k$ it is clear that $G_i \leq G$, and so (ρ_i, W_i) is a representation of G_i. Let $U_i \leq W_i$ be a (G_i, ρ_i)-covering space of W_i with codimension $h_{G_i,\rho_i}(W_i)$, and let $U = \bigoplus_{i=1}^{k} U_i \leq V$. It is easy to see that U has codimension $\sum_{i=1}^{k} h_{G_i,\rho_i}(W_i)$. It remains to show that U is a (G, ρ)-covering subspace of V.

Let $v \in V = \bigoplus_{i=1}^{k} W_i$, so for $i = 1, ..., k$ let $w_{i,0} \in W_i$ be the unique elements such that $v = \sum_{i=1}^{k} w_{i,0}$. As U_1 is a (G_1, ρ_1)-covering space of W_1, there exists $g_1 \in G_1$ such that $g_1(w_{1,0}) \in U_1$. Let $v_1 = g_1(v)$, and for $i = 1, ..., k$ let $w_{i,1} \in W_i$ be the unique elements such that $v_1 = \sum_{i=1}^{k} w_{i,1}$. In particular $w_{1,1} = g_1(w_{1,0}) \in U_1$.

Suppose for some $1 \leq t < k$ we have found for $j = 1, ..., t$ a $g_j \in G_j$ and arrived at $g_t(g_{t-1}(...(g_1(v)))) = v_t = \sum_{i=1}^{k} w_{i,t}$ such that $w_{i,t} = w_{i,i} \in U_i$ for $i = 1, ..., t$. Then, since U_{t+1} is a (G_{t+1}, ρ_{t+1})-covering space of W_{t+1}, there exists $g_{t+1} \in G_{t+1}$ such that $g_{t+1}(w_{t+1,t}) \in U_{t+1}$. Furthermore, as $G_{t+1} = \bigcap_{j=1}^{t} K_j$ we see $g_{t+1}(w_{i,t}) = w_{i,t} = w_{i,i}$ for $i = 1, ..., t$. We may therefore take $v_{t+1} = g_{t+1}(v_t)$ and proceed inductively.

This process terminates at $g_k(g_{k-1}(...(g_1(v)))) = v_k = \sum_{i=1}^{k} w_{i,i}$ such that $w_{i,i} \in U_i$ for $i = 1, ..., k$ i.e., $v_k \in U$ and therefore U is a (G, ρ)-covering space for V. Hence $h_{G,\rho}(V) \geq \text{codim}(U) = \sum_{i=1}^{k} h_{G_i,\rho_i}(W_i)$. □

Clearly the above bound can be optimised over the choice of decomposition into invariant subspaces and the order of those subspaces. It is not immediately clear how to do this.

We also have the following easy general upper bound.

Lemma 2.4.3. *Let G be a group, let (ρ, V) be a representation of G, and suppose that*

$V = \bigoplus_i W_i$, where each W_i is a G-invariant subspace of V. Then

$$h_{G,\rho}(V) \leq \sum_i h_{G,\rho}(W_i).$$

Proof. Let U be a (G,ρ)-covering subspace of V with codimension $h_{G,\rho}(V)$, and let $W_i^* = W_i \cap U$ for each i. Then W_i^* is clearly a (G,ρ)-covering subspace of W_i, and therefore $\operatorname{codim}(W_i^*) \leq h_{G,\rho}(W_i)$.

We have $\bigoplus_i W_i^* \leq U \leq V$, and therefore

$$h_{G,\rho}(V) = \operatorname{codim}(U) \leq \operatorname{codim}\left(\bigoplus_i W_i^*\right) = \sum_i \operatorname{codim}(W_i^*) \leq \sum_i h_{G,\rho}(W_i),$$

proving the lemma. □

The following is an immediate corollary of (2.6) and Lemmas 2.4.1 and 2.4.3.

Corollary 2.4.1. *Let q be a prime power, let G be a finite group with order coprime to q, let V be a finite-dimensional vector space over $\mathbb{F}_q$, and let (ρ, V) be a representation of G. Let $V = \bigoplus_i W_i$ be a decomposition of V into subrepresentations. Then*

$$\max_i\{h_{G,\rho}(W_i)\} \leq h_{G,\rho}(V) \leq \min\{\sum_i h_{G,\rho}(W_i), \lfloor \log_q(|G|)\rfloor\}. \tag{2.7}$$

We remark that, under the hypotheses of Corollary 2.4.1, Maschke's theorem guarantees the existence of a decomposition of $V = \bigoplus_i W_i$ where each W_i is an irreducible subrepresentation of V.

Theorem 2.3.1 implies that for the rotation action ρ_σ of C_{q^d-1} on $\mathbb{F}_q^{q^d-1}$, the lower bound in (2.7) is tight, as is the $\lfloor \log_q(|G|)\rfloor$ upper bound, for all $d \in \mathbb{N}$. We remark that there are infinitely many (nontrivial) cases where the sum bound is tight, and distinct from both the lower bound and the $\log_q(|G|)$ bound.

Indeed, let $m_1, ..., m_k \in \mathbb{N}$ be chosen so that for distinct $i, j \in [k]$ we have $\gcd(2^{m_i} -$

$1, 2^{m_j} - 1) = 1$ (for infinitely many examples of such, take the m_i's to be distinct primes, since if $d \mid 2^a - 1$ and $d \mid 2^b - 1$ then $d \mid 2^{\gcd(a,b)} - 1$). For each $i \in [k]$ let $n_i = 2^{m_i} - 1$, let $N = \prod_{i=1}^{k} n_i$, let $\omega_i \in \overline{\mathbb{F}}_2$ be a primitive n_i^{th} root of unity and let $f_i(X) \in \mathbb{F}_2[X]$ be the minimal polynomial of ω_i. As in the proof of Theorem 2.3.1 we know that $f_i(X)$ is irreducible, is a factor of $X^{n_i} - 1$ and has degree m_i. Let V be the vector space:

$$V = \bigoplus_{i=1}^{k} \frac{\mathbb{F}_2[X]}{\langle f_i(X) \rangle}.$$

Now V can be identified as a multplication by X invariant subspace of $\frac{\mathbb{F}_2[X]}{\langle X^N - 1 \rangle}$ by applying the Chinese Remainder theorem to the map θ from Theorem 2.3.1. Hence V is a representation of C_N where the generator of C_N acts on ordered tuples of polynomials in V by multiplying pointwise by X. We also note that $|V| = 2^{\sum_{i=1}^{k} \deg(f_i)} = 2^{\sum_{i=1}^{k} m_i}$.

Let

$$e_i = (0, ..., 0, \underbrace{1}_{i^{\text{th}}}, 0, ..., 0) \in V$$

and let $U = \text{Span}\{e_1, ..., e_k\} \leq V$. For $A \subseteq [k]$ we see that the orbit of $\sum_{i \in A} e_i$ under pointwise multiplication by powers of X has size $\prod_{i \in A}(2^{m_i} - 1)$. Indeed, from identical reasoning to that in Theorem 2.3.1 the orbit of e_i is

$$\{0\} \times ... \times \{0\} \times \underbrace{\left(\frac{\mathbb{F}_2[X]}{\langle f_i(X) \rangle} \setminus \{0\} \right)}_{i^{\text{th}}} \times \{0\} \times .... \times \{0\},$$

which has size $2^{m_i} - 1$. Now the orbit of $\sum_{i \in A} e_i$ (which we will denote $\text{Orb}(\sum_{i \in A} e_i)$) clearly has size dividing $\prod_{i \in A}(2^{m_i} - 1)$, and if $X^r \sum_{i \in A} e_i = \sum_{i \in A} e_i \iff X^r e_i = e_i$ for all $i \in A \Rightarrow (2^{m_i} - 1) \mid r$ for all $i \in A$. Since the m_i were chosen so that the $2^{m_i} - 1$ are pairwise coprime, we see that $\prod_{i \in A}(2^{m_i} - 1) \mid r$, and so the orbit of $\sum_{i \in A} e_i$ has size $\prod_{i \in A}(2^{m_i} - 1)$ as claimed.

These orbits are also disjoint, since multiplication by X preserves the positions of

zeros in the tuple. Hence

$$\begin{aligned}
|\bigcup_{A\subseteq[k]} \mathrm{Orb}(\sum_{i\in A} e_i)| &= \sum_{A\subseteq[k]} |\mathrm{Orb}(\sum_{i\in A} e_i)| \\
&= \sum_{A\subseteq[k]} \prod_{i\in A} (2^{m_i}-1) \\
&= \sum_{A\subseteq[k-1]} \left[\prod_{i\in A}(2^{m_i}-1) + (2^{m_k}-1)\prod_{i\in A}(2^{m_i}-1)\right] \\
&= 2^{m_k} \sum_{A\subseteq[k-1]} \prod_{i\in A}(2^{m_i}-1) \\
&= 2^{\sum_{i=1}^k m_i} \quad \text{(by induction on } k\text{)} \\
&= |V|,
\end{aligned}$$

and we see that every $v \in V$ lies in the orbit of some $\sum_{i\in A} e_i$, and therefore that U is a C_N-covering subspace of V. The codimension of U is $\sum_{i=1}^k (m_i - 1) = \sum_{i=1}^k h_{C_N}\left(\frac{\mathbb{F}_2[X]}{\langle f_i(X)\rangle}\right)$, showing the sum bound is tight, i.e., $h_{C_N}(V) = \sum_{i=1}^k (m_i - 1)$.

For comparison, the log bound in this case is:

$$\log_2(N) = \sum_{i=1}^k \log_2(2^{m_i}-1) = \sum_{i=1}^k m_i + \log_2\left(\prod_{i=1}^k (1-2^{-m_i})\right)$$

So if $m_i \geq 2$ for all i then $\prod_{i=1}^k (1-2^{-m_i}) \geq \left(\frac{3}{4}\right)^k$, so $\log_2(N) \geq \sum_{i=1}^k m_i - k\log_2(4/3) > h_{C_N}(V)$. It is clear that this example generalises to examples over $\mathbb{F}_q$ where q is a prime power. We also note that $h_2(N) \geq h_{C_N}(V)$, so these calculations also provide further lower bounds for h_2.

We can abstract the above argument to get the following.

Lemma 2.4.4. *Let $N \in \mathbb{N}$, let V be a vector space over field $\mathbb{F}$ such that Char$(\mathbb{F}) \nmid N$, and let (ρ, V) be a representation of the cyclic group C_N. By Maschke's theorem we may decompose $V = \bigoplus_i W_i$ into a sum of irreducible subrepresentations. Suppose that for*

$i \neq j$: if $A \subseteq W_i$ and $B \subseteq W_j$ are orbits under C_N then gcd$(|A|, |B|) = 1$.

Then:

$$\sum_i h_{C_N,\rho}(W_i) \leq h_{C_N,\rho}(V).$$

Proof. Let $N \in \mathbb{N}$ and (ρ, V) be a representation of C_N over $\mathbb{F}$ with $\mathrm{Char}(\mathbb{F}) \nmid N$. Write $V = \bigoplus_i W_i$ as a direct sum of irreducible representations. For each i take $U_i \leq W_i$ a C_N-covering subspace of W_i with codimension $h_{C_N}(W_i)$ (we have suppressed ρ in our notation for readability). Then set $U = \bigoplus_i U_i \leq V$, and note that $\mathrm{codim}(U) = \sum_i h_{C_N}(W_i)$.

We claim that U is a C_N-covering subspace of V. Indeed, let $v \in V$. As $V = \bigoplus_i W_i$ there are unique $w_i \in W_i$ such that $v = \sum_i w_i$. It immediately follows that the orbit of v under C_N is contained in $\{\sum_i x_i \mid x_i \in \mathrm{Orbit}(w_i) \subseteq W_i\}$, which has size $\prod_i |\mathrm{Orbit}(w_i)|$.

Let $g \in C_N$ be a generator for the group, so the orbit of v is $\{v, g(v), g^2(v), ..., g^{k-1}(v)\}$ where k is the smallest positive integer such that $g^k(v) = v$. Now $g^j(v) = \sum_i g^j(w_i)$, so $|\mathrm{Orbit}(w_i)| \mid k$ for each i. By assumption these orbit sizes are pairwise coprime, so $\prod_i |\mathrm{Orbit}(w_i)| \mid k$. We deduce that $\mathrm{Orbit}(v) = \{\sum_i x_i \mid x_i \in \mathrm{Orbit}(w_i) \subseteq W_i\}$.

Now, for each i we have that U_i is C_N-covering for W_i, so $\mathrm{Orbit}(w_i) \cap U_i \neq \phi$. Let $y_i \in \mathrm{Orbit}(w_i) \cap U_i$, so $\sum_i y_i \in U \cap \{\sum_i x_i | x_i \in \mathrm{Orbit}(w_i) \subseteq W_i\} = U \cap \mathrm{Orbit}(v)$. Thus U is C_N-covering for V and so $h_{C_N}(V) \geq \mathrm{codim}(U) = \sum_i h_{C_N}(W_i)$ as claimed. □

Combining this lemma with Lemma 2.4.3 we see that the conditions of Lemma 2.4.4 imply $h_{C_N}(V) = \sum_i h_{C_N}(W_i)$.

We can also use the representation theory perspective to give a slight improvement on Lemma 2.2.2.

Lemma 2.4.5. *Let q be a prime power, let $n, m \in \mathbb{N}$ and suppose m is coprime to Char$(\mathbb{F}_q)$. Then:*

$$h_q(nm) \geq h_q(n) + h_q(m).$$

Proof. Let q be a prime power, let $n, m \in \mathbb{N}$ and suppose that m is coprime to $\mathrm{Char}(\mathbb{F}_q)$. As in the proof of Theorem 2.3.1 we identify $\mathbb{F}_q^{nm}$ with $\mathbb{F}_q[X]/\langle X^{nm}-1\rangle$ and we decompose the latter into invariant subspaces as follows.

Note the factorisation

$$X^{nm}-1 = (X^n-1)(X^{(m-1)n}+X^{(m-2)n}+...+X^n+1)$$

and label these factors $f_1(X) = X^n - 1$ and $f_2(X) = X^{(m-1)n}+X^{(m-2)n}+...+X^n+1$. Define a linear map

$$\theta : \frac{\mathbb{F}_q[X]}{\langle X^{nm}-1\rangle} \to \frac{\mathbb{F}_q[X]}{\langle f_1(X)\rangle} \oplus \frac{\mathbb{F}_q[X]}{\langle f_2(X)\rangle}; \quad \theta(p(X)) = (p(X) \bmod f_1(X), p(X) \bmod f_2(X)),$$

i.e., θ is the direct sum of the natural quotient maps corresponding to the ideals generated by f_1 and f_2.

Now f_1 and f_2 are coprime: indeed, if $\alpha \in \overline{\mathbb{F}}_q$ is a common root of f_1 and f_2, then from $f_1(\alpha) = 0$ we see $\alpha^n = 1$ and substituting this into $f_2(\alpha)$ implies $m \equiv 0$ in $\mathbb{F}_q$. This contradicts the assumption that m is coprime to $\mathrm{Char}(\mathbb{F}_q)$, and so no such α exists.

It follows from the Chinese Remainder Theorem for rings that θ is a linear isomorphism. Define

$$V_1 = \left\{ p(X) \in \frac{\mathbb{F}_q[X]}{\langle X^{nm}-1\rangle} : f_2(X) \text{ divides } p(X) \right\}$$
$$V_2 = \left\{ p(X) \in \frac{\mathbb{F}_q[X]}{\langle X^{nm}-1\rangle} : f_1(X) \text{ divides } p(X) \right\}.$$

Since we have

$$V_1 = \theta^{-1}\left(\frac{\mathbb{F}_q[X]}{\langle f_1(X)\rangle} \times \{0\} \right)$$
$$V_2 = \theta^{-1}\left(\{0\} \times \frac{\mathbb{F}_q[X]}{\langle f_2(X)\rangle} \right),$$

and θ is a linear isomorphism, we have the direct sum decomposition

$$\frac{\mathbb{F}_q[X]}{\langle X^{nm}-1\rangle} = V_1 \oplus V_2,$$

where V_1 and V_2 may be viewed as copies of $\mathbb{F}_q[X]/\langle f_1(X)\rangle$ and $\mathbb{F}_q[X]/\langle f_2(X)\rangle$ respectively. Moreover, each V_i is closed under multiplication by X (i.e., under the cyclic action V_1 and V_2 are invariant subspaces).

Let $v_1(X) \in V_1 \leq \mathbb{F}_q[X]/\langle X^{nm}-1\rangle$ such that $v_1(X) \equiv 1 \bmod f_1(X)$. Then we claim the cyclic orbit of $v_1(X)$ is

$$\text{Orbit}(v_1(X)) = \{v_1(X), Xv_1(X), X^2v_1(X), ..., X^{n-1}v_1(X)\}$$

and furthermore

$$\text{Span}(\text{Orbit}(v_1)) = V_1$$

To show this, first observe that $X^r v_1(X) \neq v_1(X)$ for $1 \leq r \leq n-1$. Indeed, suppose for a contradiction there exists $r \in [n-1]$ such that multiplication by X^r fixes $v_1(X)$ in $\mathbb{F}_q[X]/\langle X^{nm}-1\rangle$. Then $X^r v_1(X) \equiv v_1(X) \bmod (X^{nm}-1)$, and therefore $X^{nm}-1$ divides $(X^r-1)v_1(X)$. It follows that X^n-1 divides $(X^r-1)v_1(X)$, but since $v_1(X) \equiv 1 \bmod (X^n-1)$ we see $(X^n-1) \mid (X^r-1)$. Hence $r=0$ or $r \geq n$, either way contradicting our assumption.

Next, we observe that $X^n v_1(X) = v_1(X)$. Indeed, as $v_1(X) \in V_1$ we have $f_2(X) \mid v_1(X)$, and so

$$(X^{nm}-1) = f_1(X)f_2(X) \mid (X^n-1)v_1(X),$$

hence $X^n v_1(X) = v_1(X)$. From these two observations it follows that the orbit of $v_1(X)$ is precisely

$$\{v_1(X), Xv_1(X), X^2v_1(X), ..., X^{n-1}v_1(X)\},$$

as claimed.

Now observe that $\theta\left(\sum_{r=0}^{n-1}\lambda_r X^r v_1(X)\right) = \left(\sum_{r=0}^{n-1}\lambda_r X^r, 0\right)$, and so θ restricted to $\mathrm{Span}(\mathrm{Orbit}(v_1))$ is a surjection onto $\mathbb{F}_q[X]/\langle f_1(X)\rangle \times \{0\}$. As θ is a linear isomorphism we deduce that $\mathrm{Span}(\mathrm{Orbit}(v_1)) = V_1$ as claimed.

We now identify θ as the following linear isomorphism θ_1 when restricted to V_1:

$$\theta_1 : V_1 \to \frac{\mathbb{F}_q[X]}{\langle X^n - 1\rangle}; \quad \theta_1\left(\sum_{r=0}^{n-1}\lambda_r X^r v_1(X)\right) = \sum_{r=0}^{n-1}\lambda_r X^r$$

and note that multiplication by X in V_1 corresponds precisely to multiplication by X in $\mathbb{F}_q[X]/\langle X^n - 1\rangle$. We may therefore take $W_1 \leq \mathbb{F}_q[X]/\langle X^n - 1\rangle$ a cyclically covering subspace of codimension $h_q(n)$, then let $U_1 = \theta_1^{-1}(W_1) \leq V_1$ and note that it is cyclically covering for V_1 and of codimension $h_q(n)$.

Let $v_2(X) \in V_2 \leq \mathbb{F}_q[X]/\langle X^{nm} - 1\rangle$ such that $v_2(X) \equiv 1 \bmod f_2(X)$. Then we define

$$O_{v_2} = \{v_2(X), Xv_2(X), X^2 v_2(X), ..., X^{(m-1)n-1}v_2(X)\},$$

and claim that

$$\mathrm{Span}(O_{v_2}) = V_2.$$

Indeed, first note that all the elements of O_{v_2} are distinct: suppose that $X^r v_2 = v_2$ for some $r < n(m-1)$. Then $X^{nm} - 1 \mid (X^r - 1)v_2(X)$ and so $f_2(X) \mid (X^r - 1)$ which implies that $r \geq n(m-1)$, a contradiction. Hence all elements of O_{v_2} are distinct.

Now observe that $\theta\left(\sum_{r=0}^{(m-1)n-1}\lambda_r X^r v_2(X)\right) = \left(0, \sum_{r=0}^{(m-1)n-1}\lambda_r X^r\right)$, an so θ restricted to $\mathrm{Span}(O_{v_2})$ is a surjection onto $\{0\} \times \mathbb{F}_q[X]/\langle f_2(X)\rangle$. As θ is a linear isomorphism we deduce that $\mathrm{Span}(O_{v_2}) = V_2$ as claimed.

Now define

$$V_2' = \mathrm{Span}\left(\{v_2(X), X^n v_2(X), X^{2n}v_2(X), ..., X^{(m-2)n}v_2(X)\}\right) \leq V_2,$$
$$B = \mathrm{Span}\left(\{X^r v_2(X) \in O_{v_2} \text{ such that } n \nmid r\}\right) \leq V_2,$$

and note firstly $V_2 = V_2' \oplus B$ and secondly both V_2' and B are invariant under multiplication by X^n. Define a linear map

$$\theta_2 : V_2' \to \frac{\mathbb{F}_q[Y]}{\langle Y^{m-1} + Y^{m-2} + ... + Y + 1\rangle}; \quad \theta_2\left(\sum_{r=0}^{m-2} \lambda_r X^{rn} v_2(X)\right) = \sum_{r=0}^{m-2} \lambda_r Y^r.$$

Note that θ_2 is a linear isomorphism and that multiplication by X^n in V_2' corresponds precisely to multiplication by Y in $\mathbb{F}_q[Y]/\left\langle Y^{m-1} + Y^{m-2} + ... + Y + 1\right\rangle$. Indeed, we first observe that

$$\begin{aligned}
& X^{(m-1)n} v_2(X) = -\sum_{r=0}^{m-2} X^{rn} v_2(X) \\
\iff & (X^{nm} - 1) \mid \left(\sum_{r=0}^{m-1} X^{rn}\right) v_2(X) \\
\Leftarrow & (X^n - 1) \mid v_2(X) \\
\Leftarrow & v_2(X) \in V_2,
\end{aligned}$$

and so

$$\begin{aligned}
\theta_2\left(X^n \sum_{r=0}^{m-2} \lambda_r X^{rn} v_2(X)\right) &= \theta_2\left(\left(\sum_{r=1}^{m-2} \lambda_{r-1} X^{rn} v_2(X)\right) + \lambda_{m-2} X^{(m-1)n} v_2(X)\right) \\
&= \theta_2\left(-\lambda_{m-2} v_2(X) + \sum_{r=1}^{m-2} (\lambda_{r-1} - \lambda_{m-2}) X^{rn} v_2(X)\right) \\
&= -\lambda_{m-2} + \sum_{r=1}^{m-2} (\lambda_{r-1} - \lambda_{m-2}) Y^r \\
&= Y \sum_{r=0}^{m-2} \lambda_r Y^r \\
&= Y \theta_2\left(\sum_{r=0}^{m-2} \lambda_r X^{rn} v_2(X)\right).
\end{aligned}$$

We may therefore take $W_2 \leq \mathbb{F}_q[Y]/\left\langle Y^{m-1} + Y^{m-2} + ... + Y + 1\right\rangle$ a cyclically cov-

ering subspace of codimension $h = h_{C_m}(\mathbb{F}_q[Y]/\left\langle Y^{m-1} + Y^{m-2} + ... + Y + 1\right\rangle)$ i.e., covering under multiplication by powers of Y. Then let $U_2 = \theta_2^{-1}(W_2) \leq V_2'$ and note that this covers under multiplication by powers of X^n and has codimension h.

Now consider

$$U = U_1 \oplus (U_2 \oplus B) \leq V_1 \oplus V_2 = \frac{\mathbb{F}_q[X]}{\langle X^{nm} - 1\rangle}.$$

We observe that U cyclically covers $\mathbb{F}_q[X]/\langle X^{nm} - 1\rangle$. Indeed, let

$$u(X) \in \mathbb{F}_q[X]/\langle X^{nm} - 1\rangle = V_1 \oplus V_2,$$

so there exist unique $u_1(X) \in V_1$ and $u_2(X) \in V_2$ such that $u(X) = u_1(X) + u_2(X)$. Since U_1 cyclically covers V_1 there exists an $r \in \{0, 1, 2, ..., n-1\}$ such that $X^r u_1(X) \in U_1$, and as V_2 is invariant under multiplication by X we have $X^r u_2(X) \in V_2 = V_2' \oplus B$. Hence there exist unique $u_2'(X) \in V_2'$ and $b(X) \in B$ such that $X^r u_2(X) = u_2'(X) + b(X)$. Since U_2 covers V_2' under multiplication by powers of X^n there exists an $s \in \{0, 1, 2, ..., m-1\}$ such that $X^{sn} u_2'(X) \in U_2$, and as B is invariant under multiplication by X^n we have $X^{sn} b(X) \in B$.

Therefore

$$\begin{aligned} X^{sn+r}u(X) &= X^{sn+r}u_1(X) + X^{sn+r}u_2(X)) \\ &= X^r X^{sn} u_1(X) + X^{sn} X^r u_2(X) \\ &= X^r u_1(X) + X^{sn}\left(u_2'(X) + b(X)\right)\right) \quad \text{(as multiplication by } X^n \text{ is trivial in } V_1) \\ &= X^r u_1(X) + X^{sn} u_2'(X) + X^{sn} b(X) \in U_1 \oplus (U_2 \oplus B) = U, \end{aligned}$$

confirming that U cyclically covers $\mathbb{F}_q[X]/\langle X^{nm} - 1\rangle$.

Now we observe that the codimension of U is $h_q(n) + h$ and so $h_q(nm) \geq h_q(n) + h$. It remains to show that $h \geq h_q(m)$.

Recall that $h = h_{C_m}(\mathbb{F}_q[Y]/\left\langle Y^{m-1} + Y^{m-2} + ... + Y + 1\right\rangle)$. Noting the factorisation

$$Y^m - 1 = (Y-1)(Y^{m-1} + Y^{m-2} + ... + Y + 1),$$

and that the factors are coprime we use the Chinese Remainder Theorem to see that

$$\psi : \frac{\mathbb{F}_q[Y]}{\langle Y^m - 1\rangle} \to \frac{\mathbb{F}_q[Y]}{\langle Y-1\rangle} \oplus \frac{\mathbb{F}_q[Y]}{\langle Y^{m-1} + Y^{m-2} + ... + Y + 1\rangle};$$
$$\psi(f(Y)) = (f(Y) \bmod (Y-1), f(Y) \bmod (Y^{m-1} + Y^{m-2} + ... + Y + 1))$$

is a linear isomorphism. Suppose that $W \leq \mathbb{F}_q[Y]/\langle Y^m - 1\rangle$ is cyclically covering, and has codimension $h_q(m)$. Let $W' = \psi(W)$, which is covering for $\frac{\mathbb{F}_q[Y]}{\langle Y-1\rangle} \oplus \frac{\mathbb{F}_q[Y]}{\left\langle \sum_{r=0}^{m-1} Y^r \right\rangle}$ under pointwise multiplication by Y. Indeed, if $(f(Y), g(Y)) \in \frac{\mathbb{F}_q[Y]}{\langle Y-1\rangle} \oplus \frac{\mathbb{F}_q[Y]}{\left\langle \sum_{r=0}^{m-1} Y^r \right\rangle}$ then let $u(Y) = \psi^{-1}((f(Y), g(Y)))$. Since W is covering for $\mathbb{F}_q[Y]/\langle Y^m - 1\rangle$, there exists an $r \in \{0, 1, 2, ..., m-1\}$ such that $Y^r u(Y) \in W$. Then $\psi(Y^r u(Y)) \in W'$, and $\psi(Y^r u(Y)) = Y^r(f(Y), g(Y))$.

We note, therefore, that $(1,0) \in W'$ since $Y(1,0) = (Y,0) = (1,0)$ and W' is covering. Hence, if we let $W'' = \{g(Y) | (0, g(Y)) \in W'\} \leq \mathbb{F}_q[Y]/\left\langle \sum_{r=0}^{m-1} Y^r \right\rangle$ then $W' = (\mathbb{F}_q[Y]/\langle Y-1\rangle) \oplus W''$. Furthermore W'' must be therefore be covering for $\mathbb{F}_q[Y]/\left\langle \sum_{r=0}^{m-1} Y^r \right\rangle$ and

$$h_q(m) = \operatorname{codim}(W) = \operatorname{codim}(W') = \operatorname{codim}(W'') \leq h_{C_n}\left(\frac{\mathbb{F}_q[Y]}{\left\langle \sum_{r=0}^{m-1} Y^r \right\rangle}\right),$$

as required.

This proves that $h_q(nm) \geq h_q(n) + h_q(m)$.

□

In a later section we will investigate the behaviour of $h_{S_n,\rho}(V)$, for various representations (ρ, V) of the symmetric group S_n.

2.5 Cases in which the covering subspaces are trivial

In this section, we demonstrate the opposite behaviour to that seen in Theorems 2.3.1 and 2.3.3 for other sequences of integers.

Theorem 2.5.1. *Let p be a prime, let q be a power of p, let $k \in \mathbb{N}$ with $k \mid q-1$, and let $d \in \mathbb{N}$. Then*

$$h_q(kp^d) = 0.$$

Equivalently, if $U \leq \mathbb{F}_q^{kp^d}$ is cyclically covering, then $U = \mathbb{F}_q^{kp^d}$.

In fact, Theorem 2.5.1 is a special case of the following result for more general representations.

Theorem 2.5.2. *Let p be a prime. Let $G = A \times B$, where A is an Abelian group of exponent k dividing $q-1$, and B is a finite p-group. Let q be a power of p, let V be a finite-dimensional vector space over $\mathbb{F}_q$ and let (ρ, V) be a representation of G. Then $h_{G,\rho}(V) = 0$.*

Theorem 2.5.2, in turn, is a consequence of the following two lemmas, which may be of independent interest.

Lemma 2.5.1. *Suppose that a finite p-group Q acts on a finite p-group P by automorphisms. If H is a subgroup of P such that $\bigcup_{g \in Q} H^g = P$, then $H = P$. (Here, as usual, H^g denotes the image of H under the automorphism defined by g.)*

Proof. The proof is by induction on $|P|$. The result is clear if $|P| = 1$, so suppose that $|P| > 1$ and that the result holds for all smaller p-groups.

Write $\Phi(P)$ for the Frattini subgroup of P, i.e., the intersection of all maximal subgroups of P. The number of subgroups of index p in P is equal to $(p^d-1)/(p-1)$, where $d \in \mathbb{N}$ is such that $|P/\Phi(P)| = p^d$. (This is well-known, and follows from the facts that a subgroup of index p in P is normal, and $\Phi(P)$ is the minimal normal subgroup of P

with elementary abelian quotient; see e.g. [47, 1.D.8]). Observe that Q acts by automorphisms on the set of all index-p subgroups of P. Since Q is a p-group, the orbit-stabilizer theorem implies that every orbit of this action has size a power of p. Since the number of index-p subgroups, $(p^d - 1)/(p - 1)$, is coprime to p, one of these orbits has size one. In other words, some index-p subgroup, P_1 say, is fixed by Q. Therefore,

$$P_1 = \bigcup_{g \in Q} (H \cap P_1)^g.$$

By the induction hypothesis, $P_1 \cap H = P_1$, so $P_1 \leq H$. Since P_1 is fixed by Q, we cannot have $H = P_1$. It follows that $H = P$, completing the induction step, proving the lemma. □

Lemma 2.5.2. *Let p be prime, and let q be a power of p. Let $G = A \times B$, where A is an Abelian group of exponent k dividing $q - 1$, and B is a finite p-group. Let G act linearly on a vector space V over $\mathbb{F}_q$, where the action by B is by automorphisms of V, and let U be a subspace of V such that the union of the images of U under G cover V. Then $U = V$.*

Proof. In the case $k = 1$, this follows from Lemma 2.5.1, applied with $Q = B$ and P the additive group of $\mathbb{F}_q$ (noting that the representation action of B on V corresponds to an action on P by automorphisms). Suppose then that $k > 1$.

Every element of A (viewed as a linear endomorphism of V) has minimum polynomial dividing $X^k - 1$, which has k distinct roots in $\mathbb{F}_q$ (since $X^k - 1$ divides $X^{q-1} - 1$, which has $q - 1$ distinct roots in $\mathbb{F}_q$), so the elements of A are all diagonalisable. Since A is Abelian, its elements can be simultaneously diagonalised, so V is the direct sum of the common eigenspaces. Since B commutes with A, it fixes each of these eigenspaces, and therefore so does G. If U contains all of the eigenspaces, then we have $U = V$, as required. Hence, we may assume that $U \cap W \subset W$ for some eigenspace W. Then the images under G of $U \cap W$ cover W. However, every element of A acts as a scalar on W,

and so fixes every subspace of W. So the images of $U \cap W$ under B cover W. The result for $k = 1$ now implies that $U \cap W = W$, contrary to our assumption. □

Proof of Theorem 2.5.2. Let p be a prime. Let $G = A \times B$, where A is an Abelian group of exponent k dividing $q - 1$, and B is a finite p-group. Let q be a power of p, let V be a finite-dimensional vector space over $\mathbb{F}_q$ and let (ρ, V) be a representation of G. Let $U \leq V$ such that $\bigcup_{g \in G} \rho(g)(U) = V$. By Lemma 2.5.2, we must have $U = V$, proving the theorem. □

Theorem 2.5.1 follows quickly from Theorem 2.5.2.

Proof of Theorem 2.5.1. If $k \mid q - 1$, then $(k, p) = 1$ so $C_{kp^d} \cong C_k \times C_{p^d}$. The group C_k has exponent k, and the group C_{p^d} is a p-group, so we can apply Theorem 2.5.2 with $V = \mathbb{F}_q^n$, yielding Theorem 2.5.1. □

We remark that Theorem 2.5.1, combined with some of our previous lemmas, determines completely the zeros of h_2.

Corollary 2.5.1. *We have $h_2(n) = 0$ if and only if $n = 2^d$ for some $d \in \mathbb{N} \cup \{0\}$, and $h_2(n) = 1$ if and only if $n = 3$.*

Proof. Applying Theorem 2.5.1 with $k = 1$ yields $h_2(2^d) = 0$ for all $d \in \mathbb{N}$. Trivially, $h_2(1) = 0$, and it is easy to see that $h_2(3) = 1$. If $n > 3$ and n is not a power of 2, then n is either divisible by 6 or by some odd number greater than 3. Let m be such a divisor. Lemma 2.2.1 implies that $h_2(m) \geq 2$ for all odd $m > 3$, and it can be checked that $h_2(6) = 2$. Hence, by Lemma 2.2.2, we have $h_2(n) \geq h_2(m) \geq 2$, proving the corollary. □

It would be interesting to determine completely, for each prime power $q > 2$, the set $\{n \in \mathbb{N} : \ h_q(n) = 0\}$. We remark that there are other zeros of h_3 besides $\{k3^d : \ k \in \{1, 2\}, \ d \in \mathbb{N}\}$ (those given by Theorem 2.5.1); for example, $h_3(4) = 0$.

We now give a second proof of Theorem 2.5.1 which is more direct.

Second proof of Theorem 2.5.1. Let p be a prime, let q be a power of p, let $k \in \mathbb{N}$ with $k \mid q-1$, and let $d \in \mathbb{N}$. Since $k \mid q-1$ and the multiplicative group of $\mathbb{F}_q$ is cyclic of order $q-1$, there exists a primitive kth root of unity in $\mathbb{F}_q$. Let $c \in \mathbb{F}_q$ be one such. Let $n = kp^d$. As in the proof of Theorem 2.3.1, we identify $\mathbb{F}_q^n$ with $\mathbb{F}_q[X]/\langle X^n - 1\rangle$, and recall that under this identification, the cyclic shift σ corresponds to multiplication by X.

Observe that $1, c, c^2, ..., c^{k-1}$ are distinct roots of $X^k - 1$ in $\mathbb{F}_q$, and so we have, in $\mathbb{F}_q[X]$,

$$X^k - 1 = \prod_{r=0}^{k-1} (X - c^r). \tag{2.8}$$

We need the following well-known fact, whose proof we include for completeness.

Claim 2.5.1. *For any $f(X) = \sum_{i=0}^{t} a_i X^i \in \mathbb{F}_q[X]$, we have $f(X)^p = \sum_{i=0}^{t} a_i^p X^{ip}$*

Proof of claim. By induction on the degree of f. When $\deg(f) = 0$, the statement of the claim holds trivially. Suppose $\deg(f) = t \geq 1$, and suppose the statement of the claim holds for all polynomials of degree less than t. Write $f(X) = aX^t + g(X)$, where $a \in \mathbb{F}_q \backslash \{0\}$ and $g(X) \in \mathbb{F}_q[X]$ with $\deg(g) < t$. Then

$$f(X)^p = (aX^t + g(X))^p = \sum_{i=0}^{p} \binom{p}{i} (aX^t)^i g(X)^{p-i} = a^p X^{tp} + g(X)^p,$$

where the final equality uses the fact that p divides $\binom{p}{i}$ whenever $1 \leq i \leq p-1$. Applying the inductive hypothesis to $g(X)$ proves the claim for f, completing the induction step, and proving the claim. □

Combining Claim 2.5.1 and (2.8), we see that, in $\mathbb{F}_q[X]$,

$$X^n - 1 = (X^k - 1)^{p^d} = \prod_{r=0}^{k-1} (X^{p^d} - c^{rp^d}). \tag{2.9}$$

Observe that the k factors in the right-hand side of (2.9) are pairwise coprime.

Hence, defining the linear map

$$\theta : \mathbb{F}_q[X]/\langle X^n - 1\rangle \to \bigoplus_{r=0}^{k-1} \mathbb{F}_q[X]/\langle X^{p^d} - c^{rp^d}\rangle; \quad \theta(p(X)) = (p(X) \bmod (X^{p^d} - c^{rp^d}))_{r=0}^{k-1},$$

the Chinese Remainder Theorem for rings implies that θ is a linear isomorphism. For $r \in \{0, 1, ..., k-1\}$, define

$$V_r := \{p(X) \in \mathbb{F}_q[X]/\langle X^n - 1\rangle : \prod_{s \neq r} (X^{p^d} - c^{sp^d}) \text{ divides } p(X)\}.$$

Note that for each $r \in \{0, 1, ..., k-1\}$, we have

$$V_r = \theta^{-1}\left(\{0\} \times ... \times \{0\} \times \frac{\mathbb{F}_q[X]}{\langle X^{p^d} - c^{rp^d}\rangle} \times \{0\} \times ... \times \{0\}\right),$$

where the zeros are in each place except for the r^{th}. Since θ is a linear isomorphism, we have a direct sum decomposition

$$\frac{\mathbb{F}_q[X]}{\langle X^n - 1\rangle} = \bigoplus_{r=0}^{k-1} V_r,$$

and V_r may be viewed as the copy of $\mathbb{F}_q[X]/\langle X^{p^d} - c^{rp^d}\rangle$ in $\mathbb{F}_q[X]/\langle X^n - 1\rangle$, for each $r \in \{0, 1, ..., k-1\}$. Observe that V_r is closed under multiplication by X.

Now suppose $U \leq \mathbb{F}_q[X]/\langle X^n - 1\rangle$ is cyclically covering, that is, $\bigcup_{t=0}^{n-1} X^t U = \mathbb{F}_q[X]/\langle X^n - 1\rangle$. For each $r \in [k-1] \cup \{0\}$, define $U_r = U \cap V_r$. Since V_r is invariant under multiplication by X, clearly U_r cyclically covers V_r (i.e., $\bigcup_{i=0}^{n-1} X^i U_r = V_r$).

We will show that

$$U_r = V_r \quad \forall r \in [k-1] \cup \{0\}. \tag{2.10}$$

(This will complete the proof of the theorem.)

To prove (2.10), fix $r \in \{0, 1, ..., k-1\}$. For each $j \in \mathbb{N} \cup \{0\}$, let f_j be the unique element of V_r such that

$$f_j \equiv (X - c^r)^j \bmod (X^{p^d} - c^{rp^d}), \quad f_j \equiv 0 \bmod (X^{p^d} - c^{sp^d}) \; \forall s \in \{0, 1, ..., k-1\} \backslash \{r\};$$

equivalently, $f_j = \theta^{-1}((0, ..., 0, (X - c^r)^j, 0..., 0))$. We now make the following claim:

Claim 2.5.2. *For any $t \in [p^d - 1]$, the entire cyclic orbit of $f_{p^d - t}$ (i.e., its orbit under multiplication by X) is contained in U_r.*

Proof of claim. We proceed by induction on t. First consider the base case, $t = 1$. Observe that, in $\mathbb{F}_q[X]$, we have

$$(X - c^r) \sum_{i=0}^{p^d - 1} c^{ri} X^{p^d - i - 1} = X^{p^d} - c^{rp^d} = (X - c^r)^{p^d},$$

with the last equality following from Claim 2.5.1. Consequently, $\sum_{i=0}^{p^d-1} c^{ri} X^{p^d-i-1} = (X - c^r)^{p^d - 1}$. Hence,

$$f_{p^d-1} \equiv \sum_{i=0}^{p^d-1} c^{ri} X^{p^d-i-1} \bmod (X^{p^d} - c^{rp^d}),$$
$$f_{p^d-1} \equiv 0 \bmod (X^{p^d} - c^{sp^d}) \; \forall s \in \{0, 1, ..., k-1\} \backslash \{r\}.$$

It follows that the cyclic orbit of f_{p^d-1} is made up of linear multiples of f_{p^d-1}. Indeed, we have

$$X \sum_{i=0}^{p^d-1} c^{ri} X^{p^d-i-1} \equiv c^r \sum_{i=0}^{p^d-1} c^{ri} X^{p^d-i-1} \bmod (X^{p^d} - c^{rp^d}),$$

so

$$Xf_{p^d-1} \equiv c^r \sum_{i=0}^{p^d-1} c^{ri} X^{p^d-i-1} \bmod (X^{p^d} - c^{rp^d}),$$

$$Xf_{p^d-1} \equiv 0 \bmod (X^{p^d} - c^{sp^d}) \ \forall s \in \{0,1,...,k-1\}\backslash\{r\},$$

and therefore $Xf_{p^d-1} = c^r f_{p^d-1}$. Hence, the cyclic orbit of f_{p^d-1} is

$$\{f_{p^d-1}, c^r f_{p^d-1}, c^{2r} f_{p^d-1}, ..., c^{(n-1)r} f_{p^d-1}\}.$$

Since U_r cyclically covers V_r, U_r has nonempty intersection with the cyclic orbit of f_{p^d-1}, say $c^{jr} f_{p^d-1} \in U_r$ for some $j \in [n-1] \cup \{0\}$. Since U_r is a subspace, it is closed under scalar multiplication, and therefore contains the entire cyclic orbit of f_{p^d-1}.

Now let $t \in [p^d - 1]$, and suppose by induction that U_r contains the entire cyclic orbit of f_{p^d-t}. Since U_r cyclically covers V_r, and $f_{p^d-t-1} \in V_r$, we have

$$U_r \cap \{f_{p^d-t-1}, Xf_{p^d-t-1}, X^2 f_{p^d-t-1}, ..., X^{n-1} f_{p^d-t-1}\} \neq \emptyset,$$

say $X^a f_{p^d-t-1} \in U_r$ for some $a \in \{0,...,n-1\}$. Then $f_{p^d-t-1} \equiv X^n f_{p^d-t-1} \in X^{n-a}U$. Note that the subspace $X^{n-a}U$ is also cyclically covering, and so contains the entire cyclic orbit of f_{p^d-t}, by the inductive hypothesis. So replacing U with $X^{n-a}U$ if necessary, we may assume that $f_{p^d-t-1} \in U_r$. Since U_r is a subspace, and $f_{p^d-t-1}, f_{p^d-t} \in U_r$, we also have $c^r f_{p^d-t-1} + f_{p^d-t} \in U_r$. Observe that

$$\begin{aligned} c^r f_{p^d-t-1} + f_{p^d-t} &\equiv c^r (X-c^r)^{p^d-t-1} + (X-c^r)^{p^d-t} \bmod (X^{p^d} - c^{rp^d}) \\ &\equiv X(X-c^r)^{p^d-t-1} \bmod (X^{p^d} - c^{rp^d}) \\ &\equiv Xf_{p^d-t-1} \bmod (X^{p^d} - c^{rp^d}) \end{aligned}$$

and

$$c^r f_{p^d-t-1} + f_{p^d-t} \equiv 0 \bmod (X^{p^d} - c^{sp^d}) \ \forall s \in \{0,1,...,k-1\}\backslash\{r\}.$$

Hence, $Xf_{p^d-t-1} = c^r f_{p^d-t-1} + f_{p^d-t} \in U_r$. Iterating this argument, we see that

$X^j f_{p^d-t-1} = c^r X^{j-1} f_{p^d-t-1} + X^{j-1} f_{p^d-t} \in U_r$ for each $j \in [n-1] \cup \{0\}$. Hence,

$$\{f_{p^d-t-1}, X f_{p^d-t-1}, X^2 f_{p^d-t-1}, ..., X^{n-1} f_{p^d-t-1}\} \subset U_r,$$

i.e., the entire cyclic orbit of f_{p^d-t-1} is contained in U_r. This completes the inductive step, proving the claim. □

The $t = p^d$ case of Claim 2.5.2 implies that $\{f_0, X f_0, X^2 f_0, ..., X^{n-1} f_0\} \subset U_r$. But

$$f_0 \equiv 1 \bmod (X^{p^d} - c^{rp^d}), \quad f_0 \equiv 0 \bmod (X^{p^d} - c^{sp^d}) \ \forall s \in \{0, 1, ..., k-1\} \backslash \{r\},$$

and therefore $\{f_0, X f_0, X^2 f_0, ..., X^{n-1} f_0\}$ spans V_r. Since U_r is a subspace, it follows that $U_r = V_r$, proving (2.10).

Since $U_r = V_r$ for each $r \in \{0, 1, ..., k-1\}$, we have

$$\mathbb{F}_q[X]/\langle X^n - 1\rangle = \bigoplus_{r=0}^{k-1} V_r \subseteq U,$$

and so $U = \mathbb{F}_q[X]/\langle X^n - 1\rangle$. It follows that $h_q(kp^d) = 0$, proving the theorem.

□

2.6 Symmetrically covering subspaces

Fix a prime power q and for $n \in \mathbb{N}$ let S_n denote the symmetric group of degree n, i.e., the group of permutations of $[n]$. We define the representation:

$$\rho : S_n \times \mathbb{F}_q^n \to \mathbb{F}_q^n; \quad \left(\sigma, \sum_{i=1}^n x_i e_i\right) \mapsto \sum_{i=1}^n x_i e_{\sigma(i)} \tag{2.11}$$

That is each $\sigma \in S_n$ is considered as a linear endomorphism on $\mathbb{F}_q^n$ given by $\sigma\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i e_{\sigma(i)}$ (i.e., S_n acts on the vector space $\mathbb{F}_q^n$ by permuting the standard basis). Given a subspace $U \leq \mathbb{F}_q^n$ we call $\{\sigma(U) : \sigma \in S_n\}$ the family of *symmetric shifts* of U. We then

say that U is *symmetrically covering* if $\bigcup_{\sigma \in S_n} \sigma(U) = \mathbb{F}_q^n$. We define $h_q^{\mathrm{sym}}(n)$ to be the maximum possible codimension of a symmetrically covering subspace of $\mathbb{F}_q^n$. In the following section we investigate the behaviour of the function $h_q^{\mathrm{sym}} : \mathbb{N} \to \mathbb{N}$.

The methods used in this section are quite different to those in earlier sections, which is unsurprising and reflects the differences between symmetric and cyclic groups. We will also make use of our earlier results on general representations of groups.

In this section, if $v \in \mathbb{F}_q^n$, let us write $v(j)$ for the jth component of v (i.e., $v = \sum_{i=1}^n v(i)e_i$ with respect to the standard basis).

Our main result in this section shows that for each sufficiently large prime p and $n = n(p)$ growing sufficiently quickly $h_p^{\mathrm{sym}}(n) = (1 - o(1))n$ where the $o(1)$ term tends to 0 as $p \to \infty$. This means the trivial bound $h_p^{\mathrm{sym}}(n) \leq n$ is asymptotically tight. We also investigate the existence of non-trivial symmetric covers, i.e., for each prime power q we lower bound the smallest n for which $h_q^{\mathrm{sym}}(n) > 0$.

2.6.1 Simple bounds for symmetrically covering subspaces

We first note that the representation theoretical generalisation of Lemma 2.2.3 implies that $h_q^{\mathrm{sym}}(n) \leq \log_q(n!)$, and it follows that $n! < q$ implies $h_q^{\mathrm{sym}}(n) = 0$ (i.e., the only symmetrically covering subspace is trivial: the whole space).

We also observe that the sequence $h_q^{\mathrm{sym}}(n)$, for $n \in \mathbb{N}$, is weakly monotonically increasing. This behaviour is different to that of $h_q(n)$ which, for example, satisfies $h_q(q^d) = 0$ and $h_q(q^d - 1) = d - 1$ for all $d \in \mathbb{N}$, so h_q is not monotonic.

Lemma 2.6.1. *Let q be a prime power and $n \leq m$ positive integers. Then*

$$h_q^{sym}(n) \leq h_q^{sym}(m).$$

Proof. Let q be a prime power. Let $n \leq m \in \mathbb{N}$. Let $U \leq \mathbb{F}_q^n$ be a symmetrically covering subspace of $\mathbb{F}_q^n$ with $\mathrm{codim}(U) = h_q^{\mathrm{sym}}(n)$. Let $k = h_q^{\mathrm{sym}}(n)$. Let $\{u_1, u_2, ..., u_{n-k}\}$ be a

basis for U. For each $i \in [n-k]$, let

$$v_i = (u_i(1), ..., u_i(n), \underbrace{0, 0, ..., 0}_{m-n}) \in \mathbb{F}_q^m$$

Let $B = \{v_1, v_2, ..., v_{n-k}\} \cup \{e_{n+1}, ..., e_m\} \subseteq \mathbb{F}_q^m$ and let $V = \text{Span}(B)$. Now $|B| = (n-k) + (m-n) = m-k$, and B is linearly independent, so $\text{codim}(V) = k$. We claim that V symmetrically covers $\mathbb{F}_q^m$. Indeed, if $x \in \mathbb{F}_q^m$ then take $\pi(x)$ to be the projection of x onto the subspace spanned by $\{e_j : j \in [n]\}$, i.e.,

$$\pi(x) := (x(1), x(2), ..., x(n), \underbrace{0, ..., 0}_{m-n}) \in \mathbb{F}_q^m$$

and let

$$\psi(x) := (x(1), x(2), ..., x(n)) \in \mathbb{F}_q^n$$

be the vector obtained from $\pi(x)$ by deleting the final $m-n$ coordinates. Since U symmetrically covers $\mathbb{F}_q^n$, there exists $\sigma \in S_n$ such that $\sigma(\psi(x)) \in U$. Viewing σ as a permutation of $[m]$ which fixes the last $m-n$ elements, it follows that $\sigma(\pi(x)) \in V$, and therefore $\sigma(x) \in V$, since B contains every unit vector e_j such that $n < j \leq m$. Hence, V is symmetrically covering, as claimed, and therefore $h_q^{\text{sym}}(m) \geq \text{codim}(V) = k = h_q^{\text{sym}}(n)$, proving the lemma. □

2.6.2 Main symmetric covering theorem and an application

We now state the main result of this section.

Theorem 2.6.1. *Let p be prime, and let $k \in \mathbb{N}$. We define*

$$T = T(k,p) = k \left\lceil (\frac{1}{\log_2(\log_2(p))})^{\frac{1}{2^{2(k+9)}}} p \right\rceil.$$

Let $n \in \mathbb{N}$. Then

$$h_p^{sym}(n) \geq \frac{k-2}{k} n - T.$$

We note that if we take p to be prime and tending to infinity, and take any $k = k(p)$ an integer valued function growing to infinity as $p \to \infty$ and $n = n(p)$ an integer valued function such that $T(k,p) = o(n)$ as $p \to \infty$, then combining Theorem 2.6.1 with $h_p^{\mathrm{sym}}(n) \leq n$ implies $h_p^{\mathrm{sym}}(n) = (1 - o(1))n$ as p tends to infinity. This theorem only applies to the case of p prime, since our proof relies on a theorem of Gowers [37], which in turn can be applied (directly) only to the case where p is prime.

However, before proving Theorem 2.6.1, we will prove the following result, which uses only purely combinatorial ideas, works for $h_q^{\mathrm{sym}}(n)$ more generally (where q is any prime power), and also captures the asymptotic behaviour of $h_q^{\mathrm{sym}}(n)$ when n is sufficiently large as a function of q (to be precise, sufficiently large here means that $n = q\log(q) \cdot f(q)$, for any function $f(q)$ which tends to infinity as $q \to \infty$).

Theorem 2.6.2. *Let q be a prime power, and let $n \in \mathbb{N}$. Then*

$$h_q^{sym}(n) \geq n - \log_{\frac{q}{q-1}}(n) = n - \frac{\log(n)}{\log(\frac{q}{q-1})}.$$

Combining this with the trivial bound $h_q^{sym}(n) \leq n$, we see that for fixed q and n tending to infinity $h_q^{sym}(n) = (1 - o(1))n$

How do Theorem 2.6.1 and Theorem 2.6.2 compare? We will demonstrate the Theorems 2.6.1 and 2.6.2 are incomparable: for many pairs (p,n) Theorem 2.6.1 is stronger than Theorem 2.6.2, and for many other pairs (p,n) Theorem 2.6.2 is stronger than Theorem 2.6.1. We restrict our attention to the case $q = p$ is prime, so that the theorems can be compared.

First we consider the regime where n is 'not growing too fast' with respect to p, where 'not growing too fast' will be determined. It can be seen from Theorem 2.6.1 that for all $\varepsilon > 0$ we may fix $k \in \mathbb{N}$ sufficiently large, such that $\frac{2}{k} \leq \frac{\varepsilon}{2}$ and take $n = n(p)$ an integer function of p such that $T = o(n)$, and we will find that, for p sufficiently large so that $T \leq \frac{\varepsilon}{2}n$, $h_p^{\mathrm{sym}}(n) \geq (1 - \varepsilon)n$. We note that this includes $n(p)$ such that $T = o(n)$ and

$n = o(p)$.

We sketch an explicit example of such, and start by letting p be prime and tending to infinity. For clarity we ignore floor and ceiling operations and define for each positive integer t the function

$$\log^{(t)}(.) := \underbrace{\log_2(\log_2(...(\log_2}_{t}(.))...)),$$

i.e., the t-fold iteration of $\log_2(.)$. Set

$$k = \log^{(7)}(p),\ T = kp\left(\frac{1}{\log^{(2)}(p)}\right)^{\frac{1}{2^{2^{k+9}}}},\ n = kp\left(\frac{\log^{(3)}(p)}{\log^{(2)}(p)}\right)^{\frac{1}{2^{2^{k+9}}}}.$$

Hence

$$2^{2^{k+9}} = \left(\log^{(5)}(p)\right)^{2^9},$$

so

$$\log_2\left(\frac{T}{n}\right) = -\frac{\log^{(4)}(p)}{\left(\log^{(5)}(p)\right)^{2^9}} \to -\infty \quad (\text{as } p \to \infty)$$

and

$$\log_2\left(\frac{n}{p}\right) = \log^{(8)}(p) + \frac{\log^{(4)}(p) - \log^{(3)}(p)}{\left(\log^{(5)}(p)\right)^{2^9}} \to -\infty \quad (\text{as } p \to \infty).$$

Hence, $T = o(n)$ and $n = o(p)$ and by Theorem 2.6.1, we have that

$$h_p^{\text{sym}}(n) \geq (\frac{k-2}{k} - \frac{T}{n}) \cdot n = (1 - o(1))n$$

as $p \to \infty$. In other words, we see from Theorem 2.6.1 that for large p the asymptotic behaviour of $h_p^{\text{sym}}(n)$ (i.e., $h_p^{\text{sym}}(n)$ essentially being equal to n) emerges before n exceeds p.

This behaviour is not captured by Theorem 2.6.2. Indeed, to analyse Theorem 2.6.2

we recall the well known fact that for all real $x > 0$ we have $\log(x) \leq x - 1$, where log is the natural logarithm. Hence $\log(\frac{p}{p-1}) \leq \frac{1}{p-1}$, and we see that the lower bound provided by Theorem 2.6.2 is at most $n-(p-1)\log(n)$. Therefore, if $n \leq p\log(p)$ this lower bound is worse than the trivial lower bound of $h_p^{\mathrm{sym}}(n) \geq 0$, and consequently worse than the bound given by Theorem 2.6.1. We can therefore see that when n is 'not growing too fast' with respect to p, then the lower bound from Theorem 2.6.1 is stronger than that from Theorem 2.6.2.

However, in the regime when n is 'growing fast' with respect to p, the lower bound from Theorem 2.6.2 overtakes the lower bound from Theorem 2.6.1. Indeed, if we optimise the lower bound from Theorem 2.6.1 under the assumption that $\left(\frac{1}{\log_2(\log_2(p))}\right)^{\frac{1}{2^{2(k+9)}}} \approx 1$ (i.e., assuming that k is growing reasonably fast with respect to p, taking $k \geq \log^{(4)}(p)$ is sufficient), and imposing $T \leq \frac{5n}{2k}$ i.e.,

$$k\left\lceil\left(\frac{1}{\log_2(\log_2(p))}\right)^{\frac{1}{2^{2(k+9)}}} p\right\rceil \leq \frac{5n}{2k},$$

then the best we can do is take $k \approx \Theta(\sqrt{n/p})$, so the lower bound becomes

$$h_p^{\mathrm{sym}}(n) \geq n - \Theta(\sqrt{n/p}),$$

which is evidently worse than $h_p^{\mathrm{sym}}(n) \geq n - p\log(n)$ once n is growing sufficiently quickly with respect to p.

While it seems likely that a better general lower bound could be obtained by combining the proofs of Theorem 2.6.1 and Theorem 2.6.2, we leave this as a problem for further research.

Proof of Theorem 2.6.2. Fix q, a prime power, and let $n \in \mathbb{N}$. Let $n_1 = \left\lceil \frac{n}{q} \right\rceil$, and for $k > 1$ we let

$$n_k = \left\lceil \frac{n - \sum_{i=1}^{k-1} n_i}{q} \right\rceil.$$

Then for all $k = 1, 2, ...$ we have that $\sum_{i=1}^{k} n_i \leq n$, and so there exists $K \in \mathbb{N}$ such that $n_k = 0$ for all $k > K$. We also note that

$$n - \sum_{i=1}^{k} n_i \leq \left(1 - \frac{1}{q}\right)^k \cdot n.$$

Indeed, this is clear for $k = 1$, and for $k > 1$ we have

$$\begin{aligned} n - \sum_{i=1}^{k} n_i &= \left(n - \sum_{i=1}^{k-1} n_i\right) - \left\lceil \frac{n - \sum_{i=1}^{k-1} n_i}{q} \right\rceil \\ &\leq \left(1 - \frac{1}{q}\right)\left(n - \sum_{i=1}^{k-1} n_i\right) \\ &\leq \left(1 - \frac{1}{q}\right)^k \cdot n. \end{aligned}$$

So, for $k \geq \log_{\frac{q}{q-1}}(n) + 1$ we have $0 \leq n - \sum_{i=1}^{k} n_i < 1$, and hence, by the integrality of all the n_i, $\sum_{i=1}^{k} n_k = n$. We deduce that $K \leq \log_{\frac{q}{q-1}}(n)$.

Now let $x = x_0 \in \mathbb{F}_q^n$. By the pigeonhole principle, at least $\left\lceil \frac{n}{q} \right\rceil = n_1$ entries of x_0 are equal. Say $\{i_{0,1}, i_{0,2}, ..., i_{0,n_1}\} \subset [n]$ such that

$$x_0(i_{0,1}) = x_0(i_{0,2}) = ... = x_0(i_{0,n_1}).$$

Take $\sigma_0 \in S_n$ such that $\sigma_0(i_{0,t}) = t$ for $t = 1, 2, ..., n_1$ and let $x_1 = \sigma_0(x_0)$ i.e., we may permute the entries of x_0 to get x_1 such that

$$x_1(1) = x_1(2) = ... = x_1(n_1),$$

and we can write

$$x_1 = (\underbrace{\mathbf{x_{1,1}}}_{n_1}, \underbrace{\mathbf{x_{1,2}}}_{n-n_1}),$$

where $\underbrace{\mathbf{x_{1,1}}}_{n_1} = (x_1(1), x_1(2), ..., x_1(n_1))$ is a vector of all equal entries and $\underbrace{\mathbf{x_{1,2}}}_{n-n_1} = (x_1(n_1 + 1), x_1(n_1 + 2), ..., x_1(n))$ are the remaining entries.

Suppose now that we have permuted the entries successively to create the sequence $x_0, x_1, ..., x_k$ such that for $1 \leq r \leq k$ we have

$$x_r = (\underbrace{\mathbf{x_{1,1}}}_{n_1}, \underbrace{\mathbf{x_{2,2}}}_{n_2}, ..., \underbrace{\mathbf{x_{r,r}}}_{n_r}, \underbrace{\mathbf{x_{r,r+1}}}_{n-\sum_{i=1}^{r} n_i})$$

where $\mathbf{x_{j,j}} = (x_j(\sum_{i=1}^{j-1} n_i + 1), x_j(\sum_{i=1}^{j-1} n_i + 2), ..., x_j(\sum_{i=1}^{j} n_i))$ is a vector of all equal entries for $j = 1, 2, ..., r$ and $\mathbf{x_{r,r+1}} = (x_r(\sum_{i=1}^{r} n_i + 1), x_r(\sum_{i=1}^{r} n_i + 2), ..., x_r(n))$.

Now, since $\mathbf{x_{k,k+1}}$ has $n - \sum_{i=1}^{k} n_i$ entries, by the pigeonhole principle at least $\left\lceil \frac{n-\sum_{i=1}^{k} n_i}{q} \right\rceil = n_{k+1}$ of these entries are equal. Say

$$\{i_{k,1}, i_{k,2}, ..., i_{k,n_{k+1}}\} \subset \left\{ \sum_{i=1}^{k} n_i + 1, \sum_{i=1}^{k} n_i + 2, ..., n \right\},$$

such that

$$x_k(i_{k,1}) = x_k(i_{k,2}) = ... = x_k(i_{k,n_{k+1}}),$$

and take $\sigma_k \in S_n$ such that $\sigma_k(i_{k,t}) = \sum_{i=1}^{k} n_i + t$ for $t = 1, 2, ..., n_{k+1}$ and $\sigma_k(i) = i$ for $i \leq \sum_{i=1}^{k} n_i$. Then take $x_{k+1} = \sigma_k(x_k)$, i.e., we may permute the entries of $\mathbf{x_{k,k+1}}$ to get

$$(\underbrace{\mathbf{x_{k+1,k+1}}}_{n_{k+1}}, \underbrace{\mathbf{x_{k+1,k+2}}}_{n-\sum_{i=1}^{k+1} n_i}),$$

where $\mathbf{x_{k+1,k+1}}$ is a length n_{k+1} vector of all equal entries and $\mathbf{x_{k+1,k+2}}$ are the remaining entries of $\mathbf{x_{k,k+1}}$. We then define

$$x_{k+1} = (\mathbf{x_{1,1}}, \mathbf{x_{2,2}}, ..., \mathbf{x_{k,k}}, \mathbf{x_{k+1,k+1}}, \mathbf{x_{k+1,k+2}}).$$

This inductive process halts after $K \leq \log_{\frac{q}{q-1}}(n)$ steps, terminating at

$$x_K = (\mathbf{x_{1,1}}, \mathbf{x_{2,2}}, ..., \mathbf{x_{K,K}}),$$

where, for $i = 1, 2, ..., K$, the block of entries $\mathbf{x_{i,i}}$ has length n_i and is a vector of all equal entries.

Let $B = \{u_j \ : \ j = 1, ..., K\}$ where

$$u_j := (\underbrace{0, ..., 0}_{\sum_{i=1}^{j-1} n_i}, \underbrace{1, ..., 1}_{n_j}, \ \underbrace{0, ..., 0}_{n-\sum_{i=1}^{j} n_i} \ ).$$

It is clear that B is linearly independent so letting $U := \text{Span}(B) \leq \mathbb{F}_q^n$ we see U has codimension $n - K$. It can easily be seen that $x_K \in U$, and we deduce that U symmetrically covers $\mathbb{F}_q^n$. Hence $h_q^{\text{sym}}(n) \geq \text{codim}(U) = n - K \geq n - \log_{\frac{q}{q-1}}(n)$. □

Our proof of Theorem 2.6.1 is similar to that for Theorem 2.6.2, but also makes use of an easy corollary of the following theorem due to Gowers [37] which extends a theorem of Roth [66] on the existence of arithmetic progressions in sets of integers. We use Gowers's Theorem as a tool to find arithmetic relationships between the entries of a vector in $\mathbb{F}_p^n$. These arithmetic relationships are then used to permute the entries in order to shift the vector into a symmetrically covering subspace of large codimension.

Theorem 2.6.3 (due to Gowers)**.** *Let $k, N \in \mathbb{N}$. Suppose $A \subseteq [N]$ with:*

$$|A| \geq (\frac{1}{\log_2(\log_2(N))})^{\frac{1}{2^{2^{(k+9)}}}} N$$

Then A contains a k-term arithmetic progression, that is $a_1, a_2, ..., a_k \in A$ satisfying $a_{i+1} - a_i = a_{j+1} - a_j$ for all $1 \leq i, j \leq k - 1$.

We now state and prove the easy corollary.

Corollary 2.6.1. *Let $k \in \mathbb{N}$ and p be prime. Let*

$$T = k \left\lceil (\frac{1}{\log_2(\log_2(p))})^{\frac{1}{2^{2^{(k+9)}}}} p \right\rceil$$

Suppose $x \in \mathbb{F}_p^T$. Then there exists $\{i_1, i_2, ..., i_k\} \subset [T]$ such that $x(i_1), x(i_2), ..., x(i_k)$ is

a, possibly trivial, arithmetic progression.

Proof. Let $k \in \mathbb{N}$ and p be prime. Let

$$T = k \left\lceil (\frac{1}{\log_2(\log_2(p))})^{\frac{1}{2^{2(k+9)}}} p \right\rceil$$

and $x \in \mathbb{F}_p^T$.

On one hand there may be $\{i_1, i_2, ..., i_k\} \subset [T]$ such that $x(i_1) = x(i_2) = ... = x(i_k)$, and so these entries form a trivial arithmetic progression.

On the other hand, if no such set exists, then for each $\alpha \in \mathbb{F}_p$ the set $\{i \in [T] : x(i) = \alpha\}$ has size less than k. Hence the $x(i)$ take at least T/k distinct values. It then follows from Theorem 2.6.3 that there is a k-term arithmetic progression amongst the $x(i)$. □

We are now ready to prove Theorem 2.6.1.

Proof of Theorem 2.6.1. Let p be prime and let $k \in \mathbb{N}$. Let

$$T = k \left\lceil (\frac{1}{\log_2(\log_2(p))})^{\frac{1}{2^{2(k+9)}}} p \right\rceil,$$

and let $n \in \mathbb{N}$. Since $h_p^{\text{sym}}(n) \geq 0$, the result of the theorem is trivially true if $n \leq T$, so without loss of generality $n \geq T$.

Let $x = x_0 \in \mathbb{F}_p^n$. Now

$$(x_0(1), x_0(2), ..., x_0(T)) \in \mathbb{F}_p^T,$$

so by Corollary 2.6.1 there exists $\{i_{0,1}, i_{0,2}, ..., i_{0,k}\} \subset [T]$ such that $x_0(i_{0,1}), x_0(i_{0,2}), ..., x_0(i_{0,k})$ is an arithmetic progression. Let σ_0 be any permutation of $[T]$ sending $i_{0,r} \to r$ for $r = 1, 2, ..., k$. Viewing σ_0 as a permutation of $[n]$ which fixes all elements greater than T, let $x_1 = \sigma_0(x_0)$, and note that $x_1(1), x_1(2), ..., x_1(k)$ is an arithmetic progression.

Suppose inductively we have permuted the entries of x to get x_m where

$$x_m((j-1)k+1), x_m((j-1)k+2), ..., x_m(jk)$$

is a k-term arithmetic progression for $j = 1, ..., m$. We now have two cases:

Case 1: If $mk + T \leq n$, then $(x_m(mk+1), x_m(mk+2), ..., x_m(mk+T)) \in \mathbb{F}_p^T$ so by Corollary 2.6.1 there exists $\{i_{m,1}, i_{m,2}, ..., i_{m,k}\} \subset \{mk+1, mk+2, ..., mk+T\}$ such that $x_m(i_{m,1}), x_m(i_{m,2}), ..., x_m(i_{m,k})$ is an arithmetic progression. Take any permutation σ_m of $\{mk+1, mk+2, ..., mk+T\}$ sending $i_{m,r} \to mk+r$ for $r = 1, 2, ..., k$. Viewing σ_m as a permutation of $[n]$ which fixes elements less than $mk+1$ or greater than $mk+T$, let $x_{m+1} = \sigma_m(x_m)$, so

$$x_{m+1}((j-1)k+1),\ x_{m+1}((j-1)k+2),\ \dots\ ,\ x_{m+1}(jk)$$

is a k-term arithmetic progression for $j = 1, ..., m+1$.

Case 2: If $mk + T > n$, then we halt the inductive process. Evidently this process halts after finitely many steps.

Once the process has halted, we have permuted x to

$$x_m = (\underbrace{\mathbf{x_{m,1}}}_{k\text{-AP}}, \underbrace{\mathbf{x_{m,2}}}_{k\text{-AP}}, ..., \underbrace{\mathbf{x_{m,m}}}_{k\text{-AP}}, \underbrace{\mathbf{x_{m,m+1}}}_{<T \text{ entries}}),$$

where $\mathbf{x_{m,j}} = (x_m((j-1)k+1), x_m((j-1)k+2), ..., x_m(jk))$ is a length k arithmetic progression for $j = 1, 2, ..., m$ and $\mathbf{x_{m,m+1}} = (x_m(mk+1), x_m(mk+2), ..., x_m(n))$.

For $j \in \{1, 2, ..., m\}$, define

$$v_j := \sum_{r=1}^{k} e_{r+(j-1)k}, \quad u_j := \sum_{r=1}^{k} (r-1)e_{r+(j-1)k}.$$

Let $B = \{v_j, u_j : j = 1, 2, ..., m\} \cup \{e_r : r = mk+1, mk+2, ..., n\}$, and let $V = \text{Span}(B) \leq$

$\mathbb{F}_p^n$. Since

$$(\underbrace{0,...,0}_{(j-1)k \text{ entries}}, a, a+d, a+2d, ..., a+(k-1)d, \underbrace{0,...,0}_{n-jk \text{ entries}}) = av_j + du_j,$$

we see that $\mathbf{x_{m,j}} \in \text{Span}(v_j, u_j)$ for $j = 1, 2, ..., m$ and

$$\mathbf{x_{m,m+1}} \in \text{Span}(\{e_r \ : \ r = mk+1, mk+2, ..., n\}).$$

Hence $x_m \in V$ and so V is symmetrically covering. Since B is linearly independent, V has dimension $2m + (n - mk) < \frac{2n}{k} + T$. Thus $\text{codim}(V) \geq \frac{k-2}{k}n - T$, proving the result. □

Theorem 2.6.1 can be applied to another collection of natural representations of symmetric groups, S_n. For $n \in \mathbb{N}$ and prime p, let $\mathbb{F}_p[S_n] = \{\sum_{\sigma \in S_n} \lambda_\sigma \sigma \mid \lambda_\sigma \in \mathbb{F}_p\}$, i.e., the $\mathbb{F}_p$-vector space with basis S_n. A natural representation of S_n on $\mathbb{F}_p[S_n]$ is given by the linear group action

$$\rho_{n,p} : S_n \times \mathbb{F}_p[S_n] \to \mathbb{F}_p[S_n]; \quad \rho_{n,p}\left(\left(\pi, \sum_{\sigma \in S_n} \lambda_\sigma \sigma\right)\right) = \sum_{\sigma \in S_n} \lambda_\sigma (\pi \circ \sigma).$$

For convenience we write $\rho_{n,p}((\pi, .))$ as $\pi(.)$ where context allows.

Theorem 2.6.4. *Let p be prime. Consider the action $\rho_{p-1,p}$. The maximum possible codimension of a subspace $U \leq \mathbb{F}_p[S_{p-1}]$ such that $\bigcup_{\sigma \in S_{p-1}} \sigma(U) = \mathbb{F}_p[S_{p-1}]$ is at least $(1 - o(1))(p-1)$ as p tends to infinity.*

Proof. Let p be prime. Consider, for $j = 1, ..., p-1$, the cosets

$$C_j = \{\sigma \in S_{p-1} \mid \sigma(p-1) = j\} = (j \ p-1)(S_{p-2}).$$

Then, let $B = \{\sum_{\sigma \in C_j} \sigma \mid j = 1, ..., p-1\}$, a linearly independent set of size $p-1$, and set $V \leq \mathbb{F}_p[S_{p-1}]$ to be the span of B. V is invariant under the action: multiplication by

$\pi \in S_{p-1}$ permutes the C_j's. Since $p = \text{char}(\mathbb{F}_p) \nmid |S_{p-1}|$, by Maschke's theorem we can find $W \leq \mathbb{F}_p[S_{p-1}]$, an invariant complement to V, such that $\mathbb{F}_p[S_{p-1}] = V \oplus W$. Hence, if $U \leq V$ covers V under the action, then $U' = U \oplus W$ covers $\mathbb{F}_p[S_{p-1}]$ under the action with $\text{codim}(U')$ in $\mathbb{F}_p[S_{p-1}]$ being equal to the $\text{codim}(U)$ in V.

Consider the linear bijection $V \to \mathbb{F}_p^{p-1}$ taking $\sum_{j=1}^{p-1} \lambda_j (\sum_{\sigma \in C_j} \sigma) \to \sum_{j=1}^{p-1} \lambda_j e_j$. Now $\pi \in S_{p-1}$ permutes the C_j by $\pi(C_j) = C_{\pi(j)}$, so it is clear that the action on V corresponds to the action $\pi(\sum_{j=1}^{p-1} \lambda_j e_j) = \sum_{j=1}^{p-1} \lambda_j e_{\pi(j)}$ on $\mathbb{F}_p^{p-1}$, that is the symmetric action we investigated in Theorem 2.6.1. The construction given there immediately gives a codimension $(1 - o(1))(p-1)$ construction for the action on V, and consequently for $\mathbb{F}_p[S_{p-1}]$. □

Combining the above theorem with the representation theoretical generalisation of Lemma 2.2.3 which shows that the maximum possible codimension of a subspace $U \leq \mathbb{F}_p[S_{p-1}]$ such that $\bigcup_{\sigma \in S_{p-1}} \sigma(U) = \mathbb{F}_p[S_{p-1}]$ is at most $\log_p(|S_{p-1}|) = \log_p(p-1)!) \leq p-1$, we find that $p-1$ is asymptotically the true value of the maximum possible codimension of a symmetric cover of $\mathbb{F}_p[S_{p-1}]$.

2.6.3 Smallest dimension for existence of non-trivial symmetric cover

We now move on to investigate the following question: What is the smallest n such that $h_q^{\text{sym}}(n) > 0$? We have already seen from the representation theoretical generalisation of Lemma 2.2.3 that if q is a prime power and $n \in \mathbb{N}$ such that $n! < q$ then $h_q^{\text{sym}}(n) = 0$, and from Lemma 2.6.1 that the sequence $h_q^{\text{syn}}(n)$ is monotone increasing in n. The following result gives a lower bound on the smallest n for which $h_q^{\text{sym}}(n) > 0$, in terms of q.

Theorem 2.6.5. *Let q be a prime power and suppose n is an integer such that $n > q^{\frac{2+o(1)}{\sqrt{\ln(q)}}}$. Then $h_q^{sym}(n) > 0$ (i.e., there exists a non-trivial symmetric covering subspace of $\mathbb{F}_q^n$).*

Proof. Let q be a prime power. The integer n will be determined during our construction.

Let k be an integer to be chosen later and take M to be the lowest common multiple of the set $[k]$. Set N to be the minimal integer such that $N > (M-1)+(\frac{M}{2}-1)+...+(\frac{M}{k}-1)$. Now let $n \in \mathbb{N}$ such that $\binom{\lfloor n/N \rfloor}{k} > q$. We now partition the index set $[n]$ into N consecutive blocks each of length $\lfloor n/N \rfloor$ or $\lceil n/N \rceil$, and call these blocks $B_1, B_2, ..., B_N$.

Suppose now that $x = (x(1), x(2), ..., x(n)) \in \mathbb{F}_q^n$. For each $t \in \{1, ..., N\}$, as $|B_t| \geq \lfloor n/N \rfloor$, there are $\binom{|B_t|}{k} > q$ sums $\sum_{i \in A} x(i)$, where $A \subseteq B_t$ and $|A| = k$, taking values in $\mathbb{F}_q$. By the pigeonhole principle there exist distinct sets $\{i_1, i_2, ..., i_k\}, \{j_1, j_2, ..., j_k\} \subseteq B_t$ with:

$$\sum_{r=1}^{k} x(i_r) = \sum_{r=1}^{k} x(j_r).$$

We will say such an expression has *true length* L if $|\{i_1, i_2, ..., i_k\} \cap \{j_1, j_2, ..., j_k\}| = k - L$. Let M_L be the number of blocks B_t which have an expression of true length L, so $\sum_{L=1}^{k} M_L \geq N$. Consequently, if $M_L \leq \frac{M}{L} - 1$ for all L we find $N \leq \sum_{L=1}^{k} (\frac{M}{L} - 1)$ contradicting our choice of N. Hence there exists an L such that $M_L \geq \frac{M}{L}$, that is there are at least $\frac{M}{L}$ blocks B_t each with an expression of true length L. Adding these $\frac{M}{L}$ disjoint expressions of true length L we find an expression of true length M

$$\sum_{r=1}^{M} x(i_r) = \sum_{r=1}^{M} x(j_r).$$

That is $i_1, ..., i_M, j_1, ..., j_M$ are all distinct. We may then permute the entries of x to get $y = (y(1), y(2), ..., y(n))$ with $\sum_{r=1}^{M} y(r) = \sum_{r=M+1}^{2M} y(r)$. Thus, setting $U = \{y \in \mathbb{F}_q^n$ such that $\sum_{r=1}^{M} y(r) = \sum_{r=M+1}^{2M} y(r)\}$, we see that U has codimension 1 and is symmetrically covering. Consequently $h_q^{\text{sym}}(n) \geq 1$.

It remains to bound n in terms of q. M is the lowest common multiple of $[k]$, so $M \leq k^{\pi(k)}$ where $\pi(k)$ is the number of primes $p \leq k$. The Prime Number Theorem [62, 68] shows $\pi(k) \leq (1 + o(1))\frac{k}{\ln(k)}$, so $M \leq e^{(1+o(1))k}$. We recall $N > (M-1) + (\frac{M}{2} - 1) + ... + (\frac{M}{k} - 1)$, so $N = M(1 + o(1)) \ln(k) \leq e^{(1+o(1))k} \ln(k)$. We require $\binom{\lfloor n/N \rfloor}{k} > q$, which is satisfied if $(\frac{n}{kN})^k > q \Leftrightarrow n > kNq^{1/k}$, so it is sufficient that $n > e^{(1+o(1))k} q^{1/k}$.

We minimise the right hand side of this condition with respect to k. The minimum occurs when $k = \sqrt{\ln q}$, so $n > q^{\frac{2+o(1)}{\sqrt{\ln(q)}}}$ is sufficient. □

It seems likely that this result can be improved, and it would be interesting to determine exactly the smallest value of n for which $h_q^{\mathrm{sym}}(n) > 0$.

2.7 Conclusion

For each prime power q, we have found infinitely many values of n such that $h_q(n) = \lfloor \log_q(n) \rfloor$ (these values of n forming certain geometric series with common ratio q or a power of q), and also infinitely many values of n such that $h_q(n) = 0$ (these values of n forming geometric progressions $(kp^d)_{d\in\mathbb{N}}$, where q is a power of the prime p and $k \mid q-1$). This demonstrates that the behaviour of $h_q(n)$ as a function of n is very irregular, depending heavily upon the prime factorisation of n. It would be interesting to determine more precisely the behaviour of $h_q(n)$ for n not of these forms. We remark that the original question of Cameron, as to whether $h_2(n)$ tends to infinity as n tends to infinity over odd integers n has been resolved by Aaronson, Groenland and Johnston [1] (subject to a conjecture of Artin [46] on the existence of infinitely many primes p for which 2 is a primitive root) after we made our results public. Isbell's conjecture however remains open.

We have also generalised the problem of finding cyclically covering subspaces to the analagous problem for a wide range of representations. We have proved some straightforward bounds for more general functions $h_{G,\rho}(V)$, and note that different bounds are tight for different input values. We have made some progress on determining $h_p^{\mathrm{sym}}(n)$, proving it is asymptotically equal to n when n is sufficiently large. Finally, we have proven a lower bound for the smallest value of n for which a non-trivial symmetric covering subspace of $\mathbb{F}_q^n$ exists. Interesting areas of further research would be to prove the asymptotic behaviour of $h_p^{\mathrm{sym}}(n)$ without resorting to the powerful theorem of Gowers, determining more precisely the minimum value of n for which non-trivial symmetric

covering subspaces exists, and calculating $h_{G,\rho}(V)$ for a wider range of groups G and representations (ρ, V).

2.8 Acknowledgements

The work in this chapter is based on the paper "Smallest cyclically covering subspaces of $\mathbb{F}_q^n$, and lower bounds in Isbell's conjecture" written jointly by Peter Cameron, David Ellis and the author. We would also like to thank Pablo Spiga for pointing out references [69, 70], and Alex Fink for useful comments after a seminar on an early version of the paper which became this chapter. We would also like to thank two anonymous referees for their careful reading of the paper, and for their helpful suggestions.

Chapter 3

Edge isoperimetric inequalities for powers of the hypercube

3.1 Introduction

3.1.1 Overview of isoperimetric problems

Isoperimetric questions are classical objects of study in mathematics. In general, they ask for the minimum possible 'boundary-size' of a set of a given 'size', where the exact meaning of these words varies according to the problem. A classical example of an isoperimetric problem is to minimise the perimeter among all shapes in the plane with unit area. The solution to this problem was 'known' to the Ancient Greeks, but the first rigorous proof was given by Weierstrass in a series of lectures in Berlin in the 1870s.

This ancient problem has since been extended to the isoperimetric problem for n-dimensional Euclidean space, $\mathbb{E}^n$, where the problem is to minimise the surface area among all bounded sets of a given volume. Again, the unique extremal set is a sphere of the required volume. More sophisticated extensions are arrived at by asking isoperimetric problems for subsets of curved surfaces. One example is the isoperimetric problem for the sphere $\mathbb{S}_2 := \{x \in \mathbb{R}^3 \ : \ x_1^2 + x_2^2 + x_3^2 = 1\}$, which is to minimise the perimeter

among connected subsets of $\mathbb{S}_2$ of given area (a problem which naturally extends to the isoperimetric problem for the n-sphere $\mathbb{S}_n := \{x \in \mathbb{R}^{n+1} \ : \ \sum_{i=1}^{n+1} x_i^2 = 1\}$). In this case the extremal sets are 'circular caps' i.e., subsets that are rotations of subsets of the form $\{x \in \mathbb{S}_2 \ : \ x_3 \geq z\}$. Another example is the isoperimetric problem for n-dimensional hyperbolic space $\mathbb{H}^n := \{x \in \mathbb{R}^{n+1} \ : \ x_1 > 0, \ x_1^2 - x_2^2 - x_3^2 - ... - x_{n+1}^2 = 1\}$.

In the last fifty years, there has been a great deal of interest in *discrete* isoperimetric inequalities. These deal with the boundaries of sets of vertices in graphs. There are two natural notions of boundary for graphs. Given a graph $G = (V, E)$ we have:

- *Vertex boundary:* For a subset $A \subseteq V$, the vertex boundary is defined to be

$$\partial_v(A) := \{u \in V \setminus A \ : \ uv \in E \text{ for some } v \in A\}.$$

The *vertex-isoperimetric problem for* G asks for the minimum possible size of the vertex-boundary of a m-element subset of V, for each $m \in \mathbb{N}$.

An example where the vertex isoperimetric problem has been solved is for the n-dimensional hypercube $Q_n := (\mathcal{P}([n]), \{\{A, B\} : |A \triangle B| = 1\})$, where the extremal vertex sets are isomorphic to initial segments of the *simplicial ordering*: $A <_{\text{simp}} B$ if $|A| < |B|$ or $|A| = |B|$ and A is earlier than B in the lexicographical ordering (i.e., $\max(A \setminus B) \leq \max(B \setminus A)$)). The following result is due to Harper [39], and stability for this result has been proved by Keevash and Long [53].

Theorem 3.1.1 (Harper's Theorem [39])**.** *Let $n \in \mathbb{N}$. Then for all vertex sets $\mathcal{A} \subseteq Q_n$, if $\mathcal{B} \subseteq Q_n$ such that $|\mathcal{B}| = |\mathcal{A}|$ and $\mathcal{B}$ is an initial segment of the simplicial ordering then*

$$|\partial_v(\mathcal{B})| \leq |\partial_v(\mathcal{A})|.$$

A further example is the vertex isoperimetric problem for the grid $[k]^n = \{(x_1, ..., x_n) : x_i \in [k] \ \forall i\}$ where $(x_1, ..., x_n)$ is joined to $(y_1, ..., y_n)$ if there exists an i such that $|x_i - y_i| = 1$ and $x_j = y_j$ for all $j \neq i$. Similarly to the hypercube, in this case

the extremal sets of vertices are initial segments of the *simplicial ordering* on $[k]^n$: $(x_1, ..., x_n) <_{\text{simp}} (y_1, ..., y_n)$ if either $\sum_i x_i < \sum_i y_i$ or $\sum_i x_i = \sum_i y_i$ and $x_j > y_j$ where $j = \min\{t : x_t \neq y_t\}$. The vertex isoperimetric problem for the infinite grid $\mathbb{Z}_+^n$ was proved by Wang and Wang [73], and for the finite grid $[k]^n$ a proof by compressions is given by Bollobás and Leader [15].

Theorem 3.1.2. *Let $k, n \in \mathbb{N}$. Then for all vertex sets $\mathcal{A} \subseteq [k]^n$, if $\mathcal{B} \subseteq [k]^d$ such that $|\mathcal{B}| = |\mathcal{A}|$ and $\mathcal{B}$ is an initial segment of the simplicial ordering on the grid $[k]^n$ then*

$$|\partial_v(\mathcal{B})| \leq |\partial_v(\mathcal{A})|.$$

Both of these examples have proofs using compression arguments, i.e., starting with an initial set of vertices we apply a sequence of transformations, each of which preserves the size of the set and does not increase its vertex boundary, and shifts the set structure closer to the extremal sets.

- *Edge boundary:* For a subset $A \subseteq V$, the edge boundary is defined to be

$$\partial(A) := \{uv \in E \ : \ v \in A, u \in V \setminus A\}.$$

The *edge-isoperimetric problem for G* asks for the minimum possible size of the edge-boundary of a m-element subset of V, for each $m \in \mathbb{N}$.

The edge isoperimetric inequality has also been solved for Q_n, where the extremal vertex sets are isomorphic to initial segments of the *binary* ordering, which we detail later.

Another example is the edge isoperimetric problem for the grid $[k]^n$. Unlike the previous examples of extremal sets, the extremal sets for the edge isoperimetric problem in $[k]^n$ are not nested. Extremal sets have the form $[a]^d \times [k]^{n-d}$ for some $a \in [k]$ and $0 \leq d \leq n$ or complements of such sets, optimised over these parameters. There are discrete phase transitions between these extremal families:

for example when $n = 2$, the extremal families form the sequence shown in Fig 3.1.

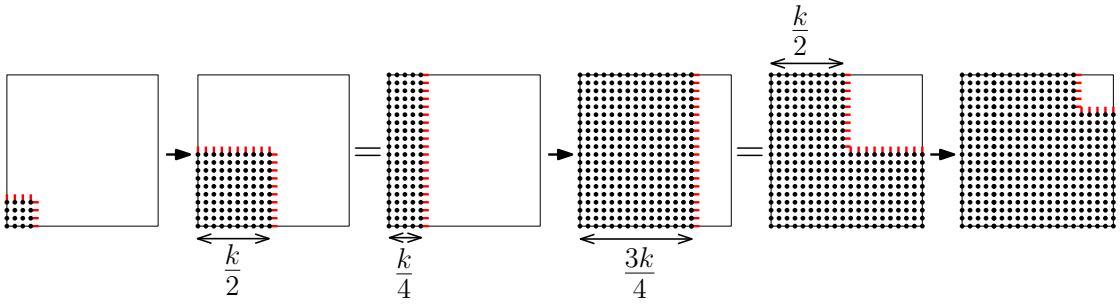


Figure 3.1: The sequence of extremal sets A for the edge isoperimetric problem in $[k]^2$, showing the phase transition at $|A| = \frac{k^2}{4}$ and $|A| = \frac{3k^2}{4}$. The edge boundaries are highlighted in red

This result was proven by Bollobás and Leader [16], by considering a related problem for subsets of the continuous cube and then recovering the discrete result.

Theorem 3.1.3. *Let $k, n \in \mathbb{N}$. Then for all vertex sets $\mathcal{A} \subseteq [k]^n$ such that $|\mathcal{A}| \leq \frac{k^n}{2}$,*

$$|\partial\mathcal{A}| \geq \min\left\{ d|\mathcal{A}|^{1-1/d}k^{n/d-1} \ : \ d = 1, 2, ..., n \right\}.$$

One final example is the edge isoperimetric problem for antipodal set systems, i.e., families $\mathcal{A} \subseteq \mathcal{P}([n])$ that are closed under taking complements. This is the natural discrete analogue of the isoperimetric problem for antipodal subsets of the n-dimensional sphere $\mathbb{S}_n$, which is a well-known unsolved problem. The extremal antipodal set systems are unions of antipodal subcubes (initial segments of the binary ordering), and this was shown by Ellis and Leader [27].

The results proved in this chapter are edge isoperimetric inequalities, i.e., for a particular graph G we will find lower bounds for $\min\{|\partial A| : \ A \subset V(G), \ |A| = m\}$, for each integer m.

3.1.2 Edge isoperimetric results for the hypercube

If $G = (V, E)$ is a graph and $A \subset V$, we write $e_G(A)$ for the number of edges of G induced by A, i.e., the number of edges of G that join two vertices in A. We remark

that if G is a regular graph, then the edge isoperimetric problem for G is equivalent to finding the maximum possible number of edges induced by a set of given size. Indeed, if G is a d-regular graph, then

$$2e_G(A) + |\partial A| = d|A| \tag{3.1}$$

for all $A \subset V$.

An important example of a discrete isoperimetric problem is the edge isoperimetric problem for the Hamming graph Q_n of the n-dimensional hypercube. We define Q_n to be the graph with vertex-set $\{0,1\}^n$, where two 0-1 vectors are adjacent if they differ in exactly one coordinate. This isoperimetric problem has numerous applications, both to other problems in mathematics, and in other areas such as distributed algorithms [2, 10], communication complexity [38], network science [12] and game theory [42].

The edge isoperimetric problem for Q_n has been solved by Harper [38], Lindsey [59], Bernstein [11] and Hart [42]. Let us describe the solution. The *binary ordering* on $\{0,1\}^n$ is defined by $x < y$ if and only if $\sum_{i=1}^n 2^{i-1}x_i < \sum_{i=1}^n 2^{i-1}y_i$. If $m \leq 2^n$, the *initial segment of the binary ordering on* $\{0,1\}^n$ *of size* m is simply the subset of $\{0,1\}^n$ consisting of the m smallest elements of $\{0,1\}^n$ with respect to the binary ordering. Note that if $m = 2^d$ for some $d \in \mathbb{N}$, then the initial segment of the binary ordering on $\{0,1\}^n$ of size m is the d-dimensional subcube $\{x \in \{0,1\}^n : \ x_i = 0 \ \forall i > d\}$.

Harper, Bernstein, Lindsey and Hart proved the following.

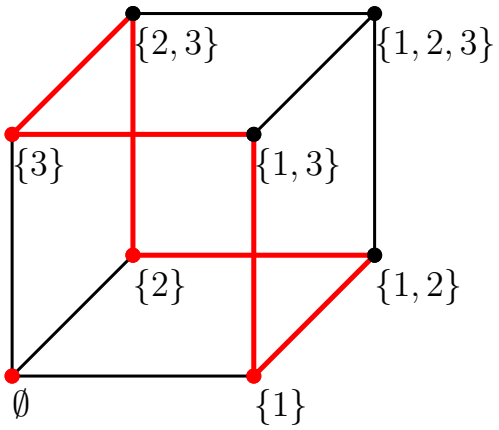
Theorem 3.1.4 (The edge isoperimetric inequality for Q_n)**.** *If $\mathcal{A} \subset \{0,1\}^n$, then $|\partial\mathcal{A}| \geq |\partial\mathcal{B}|$, where $\mathcal{B} \subset \{0,1\}^n$ is the initial segment of the binary ordering of size $|\mathcal{A}|$.*

In particular, it follows from Theorem 3.1.4 that the minimum edge-boundary of a set of size 2^d is attained by a d-dimensional subcube, for any $d \in \mathbb{N}$. As another consequence, the above theorem implies that $e_{Q_n}(\mathcal{A}) \leq \frac{1}{2}|\mathcal{A}| \log_2 |\mathcal{A}|$ for all $\mathcal{A} \subset \{0,1\}^n$.

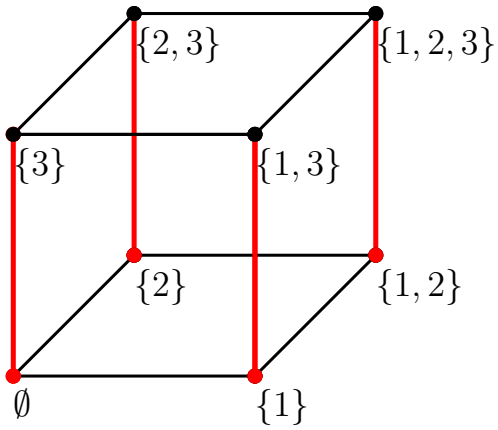
This result can be proved using induction on n and *codimension 1 compressions*: given $\mathcal{A} \subseteq \{0,1\}^n$ and $i \in [n]$ we let $\mathcal{A}_i{}^+ = \{x \in \mathcal{A} : x_i = 1\}, \mathcal{A}_i{}^- = \{x \in \mathcal{A} : x_i = 0\}$,

and then take $\mathcal{B} \subset \{0,1\}^n$ such that $|\mathcal{B}_i{}^+| = |\mathcal{A}_i{}^+|, |\mathcal{B}_i{}^-| = |\mathcal{A}_i{}^-|$ and both $\mathcal{B}_i{}^+, \mathcal{B}_i{}^-$ are initial segments of the binary ordering when the ith entry of vectors is ignored. We call $\mathcal{B}$ the codimension 1 compression of $\mathcal{A}$ in direction i.

Since $|\partial\mathcal{A}| = |\partial\mathcal{A}_i{}^+| + |\partial\mathcal{A}_i{}^-| + |\mathcal{A}_i^+ \triangle \mathcal{A}_i^-|$ (where $\partial\mathcal{A}_i{}^\pm$ and $\mathcal{A}_i^+ \triangle \mathcal{A}_i^-$ are considered within the cube obtained by ignoring the ith entry), we see by induction on n that $|\partial\mathcal{A}_i{}^\pm| \geq |\partial\mathcal{B}_i{}^\pm|$ and since $\mathcal{B}_i{}^-$ and $\mathcal{B}_i{}^+$ are nested, we have $|\mathcal{A}_i^+ \triangle \mathcal{A}_i^-| \geq |\mathcal{B}_i^+ \triangle \mathcal{B}_i^-|$. Hence $|\partial\mathcal{A}| \geq |\partial\mathcal{B}|$, so starting with an arbitrary vertex set $\mathcal{A}$ we may apply codimension 1 compressions until no further compression is possible, at which point an analysis of the resulting structure proves the result.



Edge boundary of $\{|A| \leq 1\} \subset \mathcal{P}([3])$
Example of the 'COLEX' ordering.



Edge boundary of $\mathcal{P}([2]) \subset \mathcal{P}([3])$.
Example of the 'binary' ordering.

Figure 3.2: The edge boundary (red edges) of two subsets (red vertices) of Q_3 of size 4, and for which we see the subcube ordering wins

For background on other discrete isoperimetric inequalities, we refer the reader to the surveys of Bezrukov [12] and of Leader [57].

In this chapter, we consider the edge isoperimetric problem for *powers* of the hypercube. If $r, n \in \mathbb{N}$, we let Q_n^r denote the rth power of Q_n, that is, the graph with vertex-set $\{0,1\}^n$, where two distinct 0-1 vectors are joined by an edge if they differ in at most r coordinates. Writing $[n] := \{1, 2, \ldots, n\}$, we may identify $\{0,1\}^n$ with the power-set $\mathcal{P}([n])$ via the natural bijection $x \leftrightarrow \{i \in [n] : \ x_i = 1\}$. By doing so, we may

alternatively view Q_n^r as the graph with vertex-set $\mathcal{P}([n])$, where two distinct subsets of $[n]$ are joined if their symmetric difference has size at most r. As usual, the *Hamming weight* of a vector $x \in \{0,1\}^n$ is its number of 1's; if $x, y \in \{0,1\}^n$, the *Hamming distance* between x and y is the number of coordinates on which they differ. Hence, two 0-1 vectors are adjacent in Q_n^r if and only if they are Hamming distance at most r apart.

Note that Q_n^r is a regular graph, so by (3.1), the edge isoperimetric problem for Q_n^r is equivalent to finding the maximum number of edges of Q_n^r induced by a set of given size. In other words, it is equivalent to determining

$$D(m,n,r) := \max\{e_{Q_n^r}(\mathcal{A}) : \ \mathcal{A} \subset \{0,1\}^n, \ |\mathcal{A}| = m\},$$

i.e., the maximum possible number of pairs of vectors at Hamming distance r or less, among a set of m vectors in $\{0,1\}^n$, for each $(m,n,r) \in \mathbb{N}^3$. We remark that, since Q_n^r is regular of degree $\sum_{j=1}^r \binom{n}{j}$, one has the trivial upper bound

$$D(m,n,r) \leq \tfrac{1}{2} m \sum_{j=1}^{r} \binom{n}{j} \qquad \forall m,n,r \in \mathbb{N}. \tag{3.2}$$

In the light of Theorem 3.1.4, which gives a complete answer to the isoperimetric problem for Q_n^r in the case $r = 1$, it is natural to ask whether, for each $n \geq r \geq 2$, there exists an ordering of the vertices of $\{0,1\}^n$ such that initial segments of this ordering minimize the edge-boundary in Q_n^r, over all sets of the same size. Unfortunately, this is false even for $r = 2$. Indeed, this is easy to check when $r = 2$ and $n = 4$, in which case the optimal isoperimetric sets of size 5 are precisely the Hamming balls of radius 1, whereas an optimal set of size 7 must be a 3-dimensional subcube minus a point, which contains no Hamming ball of radius 1. This indicates that the problem for $r \geq 2$ is somewhat harder than in the case $r = 1$. Still, as we shall see, reasonably good bounds can be obtained in many cases.

3.1.3 Previous results

The problem of determining (or bounding) $D(m,n,r)$ was considered by Kahn, Kalai and Linial in [51]. For half-sized sets, they solve the problem completely, proving that

$$D(2^{n-1},n,r) = 2^{n-2}\sum_{j=1}^{r}\binom{n-1}{j} \qquad \forall r,n \in \mathbb{N}. \tag{3.3}$$

(For odd r, the extremal sets for (3.3) are precisely the $(n-1)$-dimensional subcubes; for even r, the set of all vectors of even Hamming weight is also extremal.) Kahn, Kalai and Linial also observe that if $(r/n)\log(2^n/m) = o(1)$, then the trivial upper bound (3.2) is asymptotically tight, i.e.,

$$D(m,n,r) = (1-o(1))\tfrac{1}{2}m\sum_{j=1}^{r}\binom{n}{j};$$

this can be seen by considering the initial segment of the binary ordering on $\{0,1\}^n$ with size m — for example a subcube, if m is a power of 2. Finally, they observe that Kleitman's diametric theorem [56] implies that if m is 'very' small, then the 'other' trivial upper bound $D(m,n,r) \leq \binom{m}{2}$ is tight. In particular, for even values of r we know that $D(m,n,r) = \binom{m}{2}$ if and only if $m \leq \sum_{j=0}^{r/2}\binom{n}{j}$. In this case, one may consider an m-element subset of a Hamming ball of radius $r/2$, which has diameter at most r. A similar result for small sets and odd r holds as well.

It is also natural to consider the edge isoperimetric problem for the subgraph of Q_n^r induced by the binary vectors of Hamming weight k, or equivalently the graph with vertex-set $\binom{[n]}{k}$ where two k-sets are joined if their symmetric difference has size at most r. In the case $r = 2$, this graph is called the 'Kleitman-West graph', and the edge isoperimetric problem has been called the 'Kleitman-West problem' (see e.g. [40]). An elegant conjecture of Kleitman (as to the complete solution of the latter edge isoperimetric problem for all k and all vertex-set sizes) was disproved by Ahlswede and Cai [3]; only for $k \leq 2$ is a complete solution known [4, 5]. Related results have been obtained by Ahlswede and Katona [5] and Das, Gan and Sudakov [21] (Theorem 1.8 in the latter

paper implies a solution to the Kleitman-West problem for certain large values of n, for each fixed k). Harper attempted to resolve the edge isoperimetric problem in this case via a continuous relaxation [40]. Unfortunately, Harper's argument works only in certain special cases, and he later demoted his claim to a conjecture [41].

Very recently, Kirshner and Samorodnitsky [55] independently obtained isoperimetric results similar to those proved in this chapter, but using very different methods which we briefly sketch here. For any function $f : \{0,1\}^n \to \mathbb{R}$ and $p \geq 1$, as usual we define the *p-norm* of f by

$$\|f\|_p = \left(\mathbb{E}_x \left[|f(x)|^p\right]\right)^{1/p},$$

where the expectation is over a uniformly random $x \in \{0,1\}^n$. Let $H(\cdot)$ be the binary entropy function (i.e., for $q \in (0,1)$ we let $H(q) := -q\log_2(q) - (1-q)\log_2(1-q)$), and let $\psi(p,t)$ be the function on $[2,\infty) \times [0,1/2]$ defined by

$$\psi(p,t) = (p-1) + \log_2\left((1-\delta)^p + \delta^p\right) - \frac{p}{2}H(t) - pt\log_2(1-2\delta),$$

where δ is determined by $t = (\frac{1}{2} - \delta) \cdot \frac{(1-\delta)^{p-1} - \delta^{p-1}}{(1-\delta)^p + \delta^p}$. Kirshner and Samorodnitsky show that for $p \geq 2, 0 \leq s \leq \frac{n}{2}$, and for a homogeneous polynomial f of degree s on $\{0,1\}^n$ we have

$$\frac{\|f\|_p}{\|f\|_2} \leq 2^{\psi(p,s/n)\cdot\frac{n}{p}}.$$

Furthermore, they show that in a well-defined sense this inequality is 'nearly tight' if f is the Krawchouk polynomial (the Fourier transform of the characteristic function of a Hamming sphere). Kirshner and Samorodnitsky then show that these results imply for each $0 \leq s \leq n/2$ and $1 \leq r \leq 2s(1-\frac{s}{n})$ that

$$D\left(\sum_{t=0}^{s}\binom{n}{t}, n, r\right) \leq \left(\sum_{t=0}^{s}\binom{n}{t}\right) \cdot 2^{H\left(\frac{r}{2s}\right)\cdot s + H\left(\frac{r}{2(n-s)}\right)\cdot(n-s)}. \tag{3.4}$$

For odd r this upper bound is tight up to a factor of $O(\sqrt{\frac{n-s}{s}} \cdot r)$. This is compared to one of the main theorems of this chapter, Theorem 3.1.6, which is tight up to a factor

$\exp(\Theta(r))$. For fixed s, Theorem 3.1.6 is stronger for $r < \frac{1}{2}\log n$ and n sufficiently large. However, the isoperimetric bounds achieved by Kirshner and Samorodnitsky for even r, which are tight up to a factor of $O(r)$, are an improvement on the second main theorem of the chapter, Theorem 3.1.5, which is only tight up to a factor of $\exp(\Theta(r))$. Upper bound 3.4 can be applied to the Kleitman-West graph to obtain an upper bound tight up to a factor of $2e^2$. This is compared with the upper bound in Theorem 3.2.2 which is tight up to a factor of $2 + o(1)$.

3.1.4 Our results

We obtain the following bounds on $D(m, n, r)$. For brevity, we state our theorems in terms of the function $\ell = \ell(m) = \min\left\{\left\lceil \frac{2\log m}{\log n - \log\log m}\right\rceil, \lfloor \log m \rfloor\right\}$. All logs are to the base two.

Theorem 3.1.5. *Let $m, n, t \in \mathbb{N}$ with $2^t \leq m \leq 2^n$. Then*

$$D(m, n, 2t) \;\leq\; \left(\frac{8e}{t}\right)^{2t} \cdot (n \cdot \ell)^t \cdot m.$$

Theorem 3.1.6. *Let $m, n, t \in \mathbb{N}$ with $2^t \leq m \leq 2^n$. Then*

$$D(m, n, 2t+1) \;\leq\; \left(\frac{16e}{2t+1}\right)^{2t+1} \cdot (n \cdot \ell)^t \cdot m \cdot \log m.$$

We note for later use that the second term in the minimum for ℓ is the relevant one when m and n satisfy $\frac{\log m}{\log n - \log\log m} \geq \log m$, or in other words, when $m \geq 2^{n/2}$.

The two theorems above are tight up to a constant factor depending on t, viz., a factor of $\exp(\Theta(t))$; see the remark below for details. For brevity, we make no attempt to exactly optimize these constant factors. In the case $r = 2$, we prove a tighter bound (Theorem 3.2.1), which implies a new bound for the Kleitman-West problem (Theorem 3.2.2).

Determining the optimal solution to the isoperimetric problem for all vertex-set-sizes remains a challenging open problem, one which seems beyond the reach of our techniques. As mentioned above, even the restriction to k-sets and $r = 2$ is open for $k \geq 3$, that is, the Kleitman-West problem remains unsolved.

Remark 3.1.1 (Tightness)**.** *For fixed $t \in \mathbb{N}$, Theorem 3.1.5 is tight up to a factor of* $\exp(\Theta(t))$ *for some values of m, as for example can be seen by taking $\mathcal{A} = [n]^{(\leq k)}$, i.e., a Hamming ball. We suppose that $m = |\mathcal{A}| < 2^{n/2}$, then note that*

$$\begin{aligned} e_{Q_n^{2t}}(\mathcal{A}) &\geq \binom{k}{t}\binom{n-k}{t}\binom{n}{k}\frac{1}{2} \\ &\geq \left(\frac{k}{t}\right)^t \left(\frac{n-k}{t}\right)^t \binom{n}{k}\frac{1}{2} \\ &= \left(\frac{1}{t}\right)^{2t} (k(n-k))^t \binom{n}{k}\frac{1}{2} \\ &\geq \Theta\left(\left(\frac{1}{t}\right)^{2t} \cdot \left(\frac{l \cdot n}{2}\right)^t \cdot |\mathcal{A}|\right), \end{aligned}$$

where the first inequality follows from counting edges induced in the layer $[n]^{(k)}$, and the final inequality from the fact that for $m < 2^{n/2}$ we have $l = \left\lceil \frac{2 \log m}{\log n - \log \log m} \right\rceil \approx 2k$ and $n - k \approx n$.

Similarly Theorem 3.1.6 is also tight up to a factor of $\exp(\Theta(t))$*, as can be seen by taking*

$$\mathcal{A} = \{x \in [n]^{(\leq k)} : \ |x \cap \{k+1, \ldots, n\}| \leq 1\},$$

where in this case we take $k = \Theta(\log n)$.

3.1.5 Notation and Preliminaries

For subsets $\mathcal{A} \subseteq \{0,1\}^n$, we let $\mathsf{E}_{\leq r}(\mathcal{A})$ denote the set of edges in the subgraph of Q_n^r induced by vertices in $\mathcal{A}$, and we write $\mathsf{e}_{\leq r}(\mathcal{A}) := |\mathsf{E}_{\leq r}(\mathcal{A})|$. In this notation, notice that $D(m, n, r) = \max_{\mathcal{A}:|\mathcal{A}|=m} \mathsf{e}_{\leq r}(\mathcal{A})$. Abusing notation slightly, we move freely between $\{0,1\}^n$ and $\mathcal{P}([n])$ via the bijection $x \leftrightarrow \{i \in [n] : \ x_i = 1\}$. We say $\mathcal{A} \subseteq \mathcal{P}([n])$

is a *down-set* if $(x \in \mathcal{A},\ y \subseteq x) \Rightarrow y \in \mathcal{A}$. We say $\mathcal{A}$ is *left-compressed* if whenever $1 \leq i < j \leq n$ and $x \in \mathcal{A}$ with $x \cap \{i,j\} = \{j\}$, we have $(x \cup \{i\}) \setminus \{j\} \in \mathcal{A}$.

Standard compression arguments (cf. [4, 8, 41]) imply the following.

Proposition 3.1.1. *Let n, m be positive integers with $m \leq 2^n$. Among all subsets $\mathcal{A}$ of $\{0,1\}^n$ of size m, the maximum of $\mathsf{e}_{\leq r}(\mathcal{A})$ is attained where $\mathcal{A}$ is a left-compressed down-set.*

Proposition 3.1.2. *Let $\mathcal{A} \subseteq \mathcal{P}([n])$ be a down-set. For every $x \in \mathcal{A}$, we have $|x| \leq \lfloor \log |\mathcal{A}| \rfloor$.*

Proof. Since $x \in \mathcal{A}$, we also have $y \in \mathcal{A}$ for all $y \subseteq x$. The number of such y is $2^{|x|} \leq |\mathcal{A}|$. □

Remark 3.1.2. *Proposition 3.1.1 and Proposition 3.1.2 imply $\mathsf{e}_{\leq 1}(\mathcal{A}) \leq \lfloor \log |\mathcal{A}| \rfloor \cdot |\mathcal{A}|$. Indeed, for a down-set $\mathcal{A}$, we have $\mathsf{e}_{\leq 1}(\mathcal{A}) = \sum_{x \in \mathcal{A}} |x| \leq |\mathcal{A}| \cdot \lfloor \log |\mathcal{A}| \rfloor$. This approximates, up to a factor of two, the optimal bound $\mathsf{e}_{\leq 1}(\mathcal{A}) \leq (1/2) \cdot |\mathcal{A}| \cdot \lfloor \log |\mathcal{A}| \rfloor$ mentioned above [11, 38, 42, 59].*

We also make use of the following technical result to bound sums of binomial coefficients. The proof of this proposition is technical and we delay it until Section 3.5 for clarity.

Proposition 3.1.3. *For all $m \in \mathbb{N} \cup \{0\}, \lambda \in [0,1), K \in \mathbb{R}^+$ we have for $m \neq 0$*

$$\left(\frac{K}{m}\right)^m + \left(\frac{K}{m+1}\right)^{m+1} \geq \left(\frac{K}{m+\lambda}\right)^{m+\lambda},$$

and for $m = 0$

$$1 + K \geq \left(\frac{K}{\lambda}\right)^{\lambda}.$$

3.2 The distance two case

The special case of our theorem for $r = 2$ has a fairly simple proof and a tighter bound.

Theorem 3.2.1. *Let $\mathcal{A} \subset \{0,1\}^n$ satisfy $1 \leq \log|\mathcal{A}| < n$. Then*

$$\mathsf{e}_{\leq 2}(\mathcal{A}) \leq n \cdot \ell' \cdot |\mathcal{A}|,$$

where $\ell' := \min\left\{\left\lceil \frac{\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|}\right\rceil, \lfloor \log|\mathcal{A}|\rfloor\right\}$.

Using an observation of Ahlswede and Cai [4], we reduce the problem to bounding the "sum of ranks" of elements in $\mathcal{A}$. We provide a proof for completeness. Define the rank of $x \in \{0,1\}^n$ as

$$\|x\| := \sum_{j\in[n]} jx_j = \sum_{j\in x} j.$$

Lemma 3.2.1. *Let $\mathcal{A}$ be a left-compressed down-set. Then,* $\mathsf{e}_{\leq 2}(\mathcal{A}) = \sum_{x\in\mathcal{A}} \|x\|$.

Proof. Notice that $\{x,y\} \in \mathsf{E}_{\leq 2}(\mathcal{A})$ implies that either $\|y\| < \|x\|$ or vice versa. We fix $x \in \mathcal{A}$ and count y such that $\|y\| < \|x\|$. Assume that $x \neq \emptyset, \{1\}$, or the bound is trivial. We separate the cases $|y| = |x|$ and $|y| < |x|$. In the first case, we count y of the form $y = x \cup \{i\} \setminus \{j\}$, where $i < j$, $j \in x$ and $i \notin x$. The number of such y is exactly

$$\sum_{j\in x}\Big(j - 1 - |\{i \in x : \ i < j\}|\Big) \ = \ \|x\| - \binom{|x|+1}{2}.$$

For the second case, with $|y| < |x|$, there are $\binom{|x|+1}{2}$ choices for y of the form $y = x\setminus\{i,j\}$ or $y = x\setminus\{i\}$, where $i,j \in x$. As we have assumed that $\mathcal{A}$ is a left-compressed down-set, the counted pairs in both cases are in $\mathsf{E}_{\leq 2}(\mathcal{A})$. Summing over $x \in \mathcal{A}$ completes the proof. □

To obtain Theorem 3.2.1 we use the left-compressedness and down-set conditions on $\mathcal{A}$ to find an upper bound of $\|x\|$ for each $x \in \mathcal{A}$ which depends only on $|\mathcal{A}|$ and n. The

theorem then follows from summing these upper bounds over $x \in \mathcal{A}$.

Lemma 3.2.2. *Let $\mathcal{A} \subset \{0,1\}^n$ be a left-compressed down-set with $|\mathcal{A}| \geq 2$. For any $x \in \mathcal{A}$,*

$$\|x\| \leq n \cdot \ell',$$

where $\ell' = \min\left\{\left\lceil \frac{\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|}\right\rceil, \lfloor \log|\mathcal{A}|\rfloor\right\}$

Assuming this lemma, we now complete the proof of Theorem 3.2.1.

Proof of Theorem 3.2.1. Applying Proposition 3.1.1, we may assume that $\mathcal{A}$ is a left-compressed down-set. Then, Lemma 3.2.1 and Lemma 3.2.2 together imply the desired bound:

$$\mathsf{e}_{\leq 2}(\mathcal{A}) = \sum_{x \in \mathcal{A}} \|x\| \leq n \cdot \ell' \cdot |\mathcal{A}|.$$

□

Theorem 3.2.1 has the following immediate corollary for the isoperimetric problem on the Kleitman-West graph, i.e., the graph on $\binom{[n]}{k}$ where two k-element sets are joined if they have symmetric difference of size two. For $\mathcal{A} \subset \binom{[n]}{k}$, we let $e(\mathcal{A})$ denote the number of edges of this graph induced by $\mathcal{A}$.

Theorem 3.2.2. *Let $\mathcal{A} \subset \binom{[n]}{k}$ be nonempty. Then*

$$\mathsf{e}(\mathcal{A}) \leq n \cdot \ell' \cdot |\mathcal{A}|,$$

where $\ell' := \min\left\{\left\lceil \frac{\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|}\right\rceil, \lfloor \log|\mathcal{A}|\rfloor\right\}$.

We remark that Theorem 3.2.2 is tight up to a factor of $2 + o(1)$, as is evidenced by the families

$$\left\{x \in \binom{[n]}{k} : [s] \subset x\right\}$$

for $k = o(n)$ and $s \in \mathbb{N}$.

3.2.1 Proof of Lemma 3.2.2

Proposition 3.1.2 implies that $|x| \leq \lfloor \log |\mathcal{A}| \rfloor$, and thus, $\|x\| \leq n|x| \leq n\lfloor \log |\mathcal{A}| \rfloor$. Therefore, we may assume that we are in the case where $\ell' = \lceil \frac{\log |\mathcal{A}|}{\log n - \log \log |\mathcal{A}|} \rceil < \lfloor \log |\mathcal{A}| \rfloor$. We note for later use that since $\ell' = \lceil \frac{\log |\mathcal{A}|}{\log n - \log \log |\mathcal{A}|} \rceil < \lfloor \log |\mathcal{A}| \rfloor$, we have

$$2 < \frac{n}{\log |\mathcal{A}|}. \tag{3.5}$$

We use the fact that $\mathcal{A}$ is a left-compressed down-set to lower bound the number of $y \in \mathcal{A}$ that are guaranteed in $\mathcal{A}$ by the existence of $x \in \mathcal{A}$. To this end, define $\beta' := \left\lfloor \frac{n\ell'}{\log |\mathcal{A}|} \right\rfloor$, and let $x = x' \cup x''$, where $x' \subseteq \{1, \ldots, \beta'\}$ and $x'' \subseteq \{\beta' + 1, \ldots, n\}$ correspond to the integers in x with values at most β' and at least $\beta' + 1$, respectively (so that $|x| = |x'| + |x''|$). We will show that

$$\|x\| \leq \beta'|x'| + n|x''| \leq n\ell'.$$

Notice that if $|x''| = 0$, then $\|x\| = \beta'|x'| = \beta'|x| \leq \beta' \log |\mathcal{A}| \leq n\ell'$, where the inequalities use Proposition 3.1.2 and the definition of β'. Thus, we may assume that $|x'| \leq |x| - 1$ and $|x''| \geq 1$.

Consider $y \in \{0,1\}^n$ of the form $y = y' \cup y''$, where $y' \subseteq x'$, $y'' \subseteq ([\beta'] \setminus x') \cup x''$, and $|y''| \leq |x''|$. We claim every y of this form is in $\mathcal{A}$. Indeed, this follows directly from the left-compressed down-set assumption. To count such $y \in \mathcal{A}$, first define $\varepsilon_x \in [0,1)$ as the real number satisfying $2^{|x'|} = |\mathcal{A}|^{\varepsilon_x}$. We will show $|x''| \leq (1 - \varepsilon_x)\ell'$. Clearly, there are $2^{|x'|} = |\mathcal{A}|^{\varepsilon_x}$ choices for $y' \subseteq x'$ and

$$\#\text{ of choices for } y'' = \sum_{j=0}^{|x''|} \binom{\beta' + |x''| - |x'|}{j},$$

where the j^{th} term counts y'' with $|y''| = j$. Since the choice of y' is independent of y'', we know that the sum above must be at most $|\mathcal{A}|^{1-\varepsilon_x}$, otherwise we would have

guaranteed more than $|\mathcal{A}|$ distinct y in $\mathcal{A}$.

Aiming for a contradiction, we suppose that $|x''| \geq \lceil (1-\varepsilon_x)\ell' \rceil$ and $\varepsilon_x \leq 1 - 1/\ell'$. It is a standard fact that for $a, b \in \mathbb{N}$ where $b \geq 1$ we have $\binom{a}{b} \geq \left(\frac{a}{b}\right)^b$. This fact and the assumption $|x''| \geq \lceil (1-\varepsilon_x)\ell' \rceil$ imply the lower bound

$$\sum_{j=0}^{|x''|} \binom{\beta' + |x''| - |x'|}{j} \geq \left(\frac{\beta' + |x''| - |x'|}{\lceil (1-\varepsilon_x)\ell' \rceil}\right)^{\lceil (1-\varepsilon_x)\ell' \rceil} + \left(\frac{\beta' + |x''| - |x'|}{\lceil (1-\varepsilon_x)\ell' \rceil - 1}\right)^{\lceil (1-\varepsilon_x)\ell' \rceil - 1} \tag{3.6}$$

$$\geq \left(\frac{\beta' + |x''| - |x'|}{(1-\varepsilon_x)\ell'}\right)^{(1-\varepsilon_x)\ell'}, \tag{3.7}$$

where the final inequality follows by applying Proposition 3.1.3.

We note that if $a > 2$ then $\frac{\lfloor a \rfloor}{\log a} \geq \frac{2}{\log(3)}$. Using our observation in equation (3.5) we apply this fact to the definition of β' to see

$$\beta' = \left\lfloor \frac{n\ell'}{\log |\mathcal{A}|} \right\rfloor \geq \left\lfloor \frac{n}{\log |\mathcal{A}|} \right\rfloor \left\lceil \frac{\log |\mathcal{A}|}{\log n - \log\log |\mathcal{A}|} \right\rceil \geq \frac{2}{\log(3)} \log |\mathcal{A}|. \tag{3.8}$$

Observe that (3.8) and the fact $|x'| = \varepsilon_x \log |\mathcal{A}|$ together imply $\beta' - |x'| \geq (1 - \frac{\log 3}{2}\varepsilon_x)\beta'$.

We now split into the following cases:

(1) $|x'| \geq 4$,

(2) $2 \leq |x'| \leq 3$,

(3) $|x'| \leq 1$.

Case (1): $|x'| \geq 4$**.** We note that $|x'| \geq 4$ is equivalent to $\varepsilon_x \log |\mathcal{A}| \geq 4$ and this implies $\varepsilon_x > \frac{\log 3}{(2-\log 3)\log |\mathcal{A}|}$, which after rearranging is equivalent to $\frac{2-\log 3}{2}\varepsilon_x > \frac{\log 3}{2\log |\mathcal{A}|}$. Using inequality (3.8), and that $1/(1-\varepsilon_x) \geq 1$, we see $\frac{2-\log 3}{2(1-\varepsilon_x)}\varepsilon_x > \frac{1}{\beta'}$. Now, by the definition of β', the right hand side of this inequality trivially satisfies

$$\frac{1}{\beta'} \geq \frac{\frac{n\ell'}{\log|\mathcal{A}|} - \beta'}{\beta'}, \tag{3.9}$$

so rearranging we see that

$$\left(\frac{(1-\frac{\log 3}{2}\varepsilon_x)\beta'}{(1-\varepsilon_x)\ell'}\right) = \left(1 + \frac{2-\log 3}{2(1-\varepsilon_x)}\varepsilon_x\right)\frac{\beta'}{\ell'} > \frac{n}{\log|\mathcal{A}|}.$$

Using our observation that $\beta' - |x'| \geq (1-\frac{\log 3}{2}\varepsilon_x)\beta'$ we arrive at

$$\frac{\beta' + |x''| - |x'|}{(1-\varepsilon_x)\ell'} > \frac{n}{\log|\mathcal{A}|}.$$

Substituting this into the lower bound (3.7) we see

$$\sum_{j=0}^{|x''|}\binom{\beta' + |x''| - |x'|}{j} > \left(\frac{n}{\log|\mathcal{A}|}\right)^{(1-\varepsilon_x)\ell'} \geq |\mathcal{A}|^{1-\varepsilon_x},$$

giving the required contradiction.

Case (2): $2 \leq |x'| \leq 3$**.** As $|x'| \leq 3$ we have $|x''| \geq 1 \geq |x'|/3$, and so

$$\beta' + |x''| - |x'| \geq \beta' - 2|x'|/3.$$

We combine this with fact (3.8) to get $\beta' + |x''| - |x'| \geq (1-\frac{\log 3}{3}\varepsilon_x)\beta'$. Therefore

$$\frac{\beta' + |x''| - |x'|}{(1-\varepsilon_x)\ell'} \geq \left(\frac{1-\frac{\log 3}{3}\varepsilon_x}{1-\varepsilon_x}\right)\frac{\beta'}{\ell'} = \left(1 + \frac{3-\log 3}{3(1-\varepsilon_x)}\varepsilon_x\right)\frac{\beta'}{\ell'}. \tag{3.10}$$

Now, since $|x'| \geq 2$ is equivalent to $\varepsilon_x \log|\mathcal{A}| \geq 2$ we see $\varepsilon_x > \frac{3\log 3}{2(3-\log 3)\log|\mathcal{A}|}$ which after rearranging is equivalent to $\frac{3-\log 3}{3}\varepsilon_x > \frac{\log 3}{2\log|\mathcal{A}|}$. Using inequality (3.8), and that $1/(1-\varepsilon_x) \geq 1$, we see $\frac{3-\log 3}{3(1-\varepsilon_x)}\varepsilon_x > \frac{1}{\beta'}$. Now, as in the previous case, we appeal to equation (3.9) and rearrange to find

$$\left(1 + \frac{3-\log 3}{3(1-\varepsilon_x)}\varepsilon_x\right)\frac{\beta'}{\ell'} > \frac{n}{\log|\mathcal{A}|}.$$

Combining this with the inequality (3.10) we find again $\frac{\beta'+|x''|-|x'|}{(1-\varepsilon_x)\ell'} > \frac{n}{\log|\mathcal{A}|}$. Substituting this into the lower bound (3.7) gives the required contradiction.

Case (3): $|x'| \leq 1$. Suppose first that $|x'| = 0$, and so $\varepsilon_x = 0$. Then by assumption $|x''| \geq \lceil \ell' \rceil$. Hence

$$\sum_{j=0}^{|x''|} \binom{\beta' + |x''| - |x'|}{j} \geq \binom{\beta' + |x''|}{\ell'} + \binom{\beta' + |x''|}{\ell' - 1} = \binom{\beta' + |x''| + 1}{\ell'} \geq \left(\frac{\beta' + |x''| + 1}{\ell'} \right)^{\ell'},$$

and since $\beta' + |x''| + 1 = \left\lfloor \frac{n\ell'}{\log|\mathcal{A}|} \right\rfloor + |x''| + 1 > \frac{n\ell'}{\log|\mathcal{A}|}$ we see that

$$\sum_{j=0}^{|x''|} \binom{\beta' + |x''| - |x'|}{j} > \left(\frac{n}{\log|\mathcal{A}|} \right)^{\ell'} \geq |\mathcal{A}|,$$

providing the required contradiction.

Secondly, we suppose that $|x'| = 1 \leq |x''|$. In this case, we have

$$\sum_{j=0}^{|x''|} \binom{\beta' + |x''| - |x'|}{j} \geq \left(\frac{\beta' + |x''| - |x'|}{(1-\varepsilon_x)\ell'} \right)^{(1-\varepsilon_x)\ell'} \geq \left(\frac{\beta'}{(1-\varepsilon_x)\ell'} \right)^{(1-\varepsilon_x)\ell'}.$$

Now $|x'| \geq 1$ is equivalent to $\varepsilon_x \log|\mathcal{A}| \geq 1$ which implies $\varepsilon_x > \frac{\log 3}{2\log|\mathcal{A}|}$. Using inequality (3.8) we see $\varepsilon_x > 1/\beta'$, which implies $\frac{\beta'}{1-\varepsilon_x} > \frac{n\ell'}{\log|\mathcal{A}|}$. Thus, if $|x'| = 1 \leq |x''|$ then

$$\sum_{j=0}^{|x''|} \binom{\beta' + |x''| - |x'|}{j} > \left(\frac{n}{\log|\mathcal{A}|} \right)^{(1-\varepsilon_x)\ell'} = |\mathcal{A}|^{(1-\varepsilon_x)},$$

again giving a contradiction.

Since in every case we arrive at a contradiction, the assumption $|x''| \geq \lceil (1-\varepsilon_x)\ell' \rceil$ is

false and so we must have $|x''| \leq \lceil (1-\varepsilon_x)\ell' \rceil - 1 < (1-\varepsilon_x)\ell'$, and thus we conclude that

$$\|x\| \;\leq\; \beta'|x'| + n|x''| \;=\; \beta'\varepsilon_x \log|\mathcal{A}| + n|x''| \;\leq\; \varepsilon_x n\ell' + (1-\varepsilon_x)n\ell' \;=\; n\ell'.$$

3.3 The general case for even distances

In this section, we prove Theorem 3.1.5, which, using the notation defined in Section 3.1.5, is equivalent to the statement that if $\mathcal{A} \subset \{0,1\}^n$ and $t \in \mathbb{N}$ with $t \leq \log|\mathcal{A}|$, then

$$|\mathsf{E}_{\leq 2t}(\mathcal{A})| := \mathsf{e}_{\leq 2t}(\mathcal{A}) \;\leq\; \left(\frac{8e}{t}\right)^{2t} \cdot (n\cdot\ell)^t \cdot |\mathcal{A}|,$$

where

$$\ell = \ell(\mathcal{A}) := \min\left\{ \left\lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|} \right\rceil, \lfloor \log|\mathcal{A}| \rfloor \right\}.$$

We start with some more notation. For $(b,a) \in \mathbb{Z}^2_{\geq 0}$, let

$$\mathsf{E}_{(b,a)}(\mathcal{A}) := \{\{x,y\} \in \mathsf{E}_{\leq 2t}(\mathcal{A}) :\; |x \setminus y| = b,\; |y \setminus x| = a\}.$$

and define $\mathsf{e}_{(b,a)}(\mathcal{A}) := |\mathsf{E}_{(b,a)}(\mathcal{A})|$. Letting

$$\mathcal{U} = \{(b,a) \in \mathbb{Z}^2_{\geq 0} :\; b \geq a \text{ and } b + a \leq 2t\},$$

observe that we can decompose $\mathsf{E}_{\leq 2t}(\mathcal{A})$ as a disjoint union

$$\mathsf{E}_{\leq 2t}(\mathcal{A}) = \bigcup_{(b,a)\in\mathcal{U}} \mathsf{E}_{(b,a)}(\mathcal{A}),$$

and in particular, this implies,

$$\mathsf{e}_{\leq 2t}(\mathcal{A}) = \sum_{(b,a)\in\mathcal{U}} \mathsf{e}_{(b,a)}(\mathcal{A}). \tag{3.11}$$

Our strategy will be to prove upper bounds on $\mathsf{e}_{(b,a)}(\mathcal{A})$, and then combine these

to obtain the theorem. We will need a variant of the bound on $|x''|$ from the proof of Lemma 3.2.2. In what follows, we express our results using integers $\ell := \ell(\mathcal{A})$ and $\beta := \beta(\mathcal{A})$, defined in the next proposition. We also define $\ell_x := |x \cap \{\beta+1, \dots, n\}|$ for $x \in \mathcal{A}$. Intuitively, β is the threshold for 'big' elements; ℓ_x is the number of these 'big' elements; and, we will show that $\ell_x \leq \ell$.

Proposition 3.3.1. *Let $n \geq 2$ and $\mathcal{A} \subset \{0,1\}^n$ be a down-set with $|\mathcal{A}| \geq 2$. Let*

$$\ell = \min\left\{\left\lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|}\right\rceil, \lfloor \log|\mathcal{A}|\rfloor\right\}, \quad \beta = \left\lfloor \left(\frac{n}{\log|\mathcal{A}|}\right)^{1/2}\ell \right\rfloor.$$

For any $x \in \mathcal{A}$, we have the following:

(i) $|x| \cdot \beta \leq n\ell$,

(ii) $\beta^2 \leq n\ell$,

(iii) $\log^2|\mathcal{A}| \leq \frac{n}{n-1}n\ell$,

(iv) $|x|^2 \leq n\ell$,

(v) $\lfloor \log|\mathcal{A}|\rfloor \log|\mathcal{A}| \leq n\ell$.

Proof. Parts (i) and (ii) follow immediately from Proposition 3.1.2, the fact that $\log|\mathcal{A}| \leq n$ and the definitions of β and ℓ.

For part (iii), since $\log(n/\log|\mathcal{A}|) \leq n/\log|\mathcal{A}|$ we see that

$$\log^2|\mathcal{A}| \leq \frac{n\log|\mathcal{A}|}{\log(n/\log|\mathcal{A}|)}.$$

Hence, if $\ell = \left\lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|}\right\rceil$ then $\ell \geq \frac{\log|\mathcal{A}|}{\log(n/\log|\mathcal{A}|)}$ and we see the stronger statement $\log^2|\mathcal{A}| \leq n\ell$ holds, and we note this for later. On the other hand, if $\ell = \lfloor\log|\mathcal{A}|\rfloor < \left\lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|}\right\rceil$, then $n\ell \geq n(\log|\mathcal{A}| - 1)$, so it is sufficient to show $\frac{n}{n-1}n(\log|\mathcal{A}|-1) \geq \log^2|\mathcal{A}|$, which is true if and only if $\frac{n}{n-1} \leq \log|\mathcal{A}| \leq n$.

Therefore, the only remaining cases to check are when $1 \leq \log|\mathcal{A}| < \frac{n}{n-1}$. Under this assumption, $\ell = 1$ and $\log^2|\mathcal{A}| < \left(\frac{n}{n-1}\right)^2$, so as $n \geq 2$ we see that $\frac{n^2}{n-1} \geq \left(\frac{n}{n-1}\right)^2$ which in turn shows $\frac{n}{n-1}n\ell \geq \log^2|\mathcal{A}|$ as required.

For part (iv) let $x \in \mathcal{A}$. We have already seen $|x| \leq \lfloor \log|\mathcal{A}| \rfloor$ and $|x| \leq n$ is trivial. If $\ell = \left\lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|} \right\rceil$, we recall that $\log^2|\mathcal{A}| \leq n\ell$, and so $|x|^2 \leq n\ell$. On the other hand, if $\ell = \lfloor \log|\mathcal{A}| \rfloor$, then $|x|^2 \leq n\ell$. This proves (iv).

Finally, for part (v), again recall that if $\ell = \left\lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|} \right\rceil$ then $\lfloor \log|\mathcal{A}| \rfloor \log|\mathcal{A}| \leq \log^2|\mathcal{A}| \leq n\ell$ and so $\lfloor \log|\mathcal{A}| \rfloor \log|\mathcal{A}| \leq n\ell$ follows. On the other hand if $\ell = \lfloor \log|\mathcal{A}| \rfloor$, then as $\log|\mathcal{A}| \leq n$ we see $\lfloor \log|\mathcal{A}| \rfloor \log|\mathcal{A}| \leq n\ell$, completing the proof of (v). □

Lemma 3.3.1. *Let $\mathcal{A} \subset \{0,1\}^n$, $|\mathcal{A}| \geq 2$ be a left-compressed down-set. If $x \in \mathcal{A}$, then $\ell_x \leq \ell$.*

Proof. Proposition 3.1.2 implies $|x| \leq \lfloor \log|\mathcal{A}| \rfloor$, and clearly $\ell_x \leq |x|$, so we may assume that we are in the case when $\ell = \lceil \frac{2\log|\mathcal{A}|}{\log n - \log\log|\mathcal{A}|} \rceil$. Let $x = x' \cup x''$ where $x' \subseteq \{1, \ldots, \beta\}$ and $x'' \subseteq \{\beta+1, \ldots, n\}$. By definition, $|x''| = \ell_x$, and since $\mathcal{A}$ is a down-set, we know that $x'' \in \mathcal{A}$. Suppose $y \subseteq [\beta] \cup x''$ with $|y| \leq \ell_x$. As $\mathcal{A}$ is left-compressed and a down-set $y \in \mathcal{A}$. Counting such y we have

$$|\mathcal{A}| \geq \sum_{j=0}^{\ell_x} \binom{\beta + \ell_x}{j}. \tag{3.12}$$

Suppose now, for a contradiction, that $\ell_x \geq \ell + 1$. Then clearly

$$\sum_{j=0}^{\ell_x} \binom{\beta + \ell_x}{j} \geq \binom{\beta + \ell_x}{\ell} + \binom{\beta + \ell_x}{\ell - 1}.$$

Applying Proposition 3.1.3 to this inequality and combining with the lower bound (3.12) we find that

$$|\mathcal{A}| \geq \left(\frac{\beta + \ell_x}{2\log|\mathcal{A}| / \log(n/\log|\mathcal{A}|)} \right)^{2\log|\mathcal{A}|/\log(n/\log|\mathcal{A}|)}. \tag{3.13}$$

Now, since $\ell_x \geq \ell + 1$ it is clear that

$$\frac{\beta + \ell_x}{2\log|\mathcal{A}|/\log(n/\log|\mathcal{A}|)} \geq \frac{\beta + 1 + \ell}{2\log|\mathcal{A}|} \cdot \log\left(\frac{n}{\log|\mathcal{A}|}\right),$$

and so by substituting the definition of β into this inequality, we see that

$$\frac{\beta + \ell_x}{2\log|\mathcal{A}|/\log(n/\log|\mathcal{A}|)} \geq \frac{\left(\left(\frac{n}{\log|\mathcal{A}|}\right)^{1/2} + 1\right) \cdot \ell}{2\log|\mathcal{A}|} \cdot \log\left(\frac{n}{\log|\mathcal{A}|}\right) > \left(\frac{n}{\log|\mathcal{A}|}\right)^{1/2}.$$

From this, and equation (3.13) we see that

$$|\mathcal{A}| > \left(\frac{n}{\log|\mathcal{A}|}\right)^{\log|\mathcal{A}|/\log(n/\log|\mathcal{A}|)} = |\mathcal{A}|,$$

which is a contradiction. We therefore deduce that $\ell_x \leq \ell$. □

In what follows, let $\mathcal{A} \subseteq \{0,1\}^n$ be a left-compressed down-set with $1 \leq \log|\mathcal{A}| < n$. Let ℓ, β be defined as in Proposition 3.3.1. Recall that $\ell_x = |x \cap \{\beta+1, \ldots, n\}|$ equals the number of large elements in $x \in \mathcal{A}$. In our proofs, it will be helpful to order $\{0,1\}^n$ based on ℓ_x. In particular, we upper bound $\mathsf{e}_{(b,a)}(\mathcal{A})$ by partitioning the pairs $\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A})$ into two sets, based on the cases $\ell_y \leq \ell_x$ and $\ell_y > \ell_x$. By the definition of $\mathsf{E}_{(b,a)}(\mathcal{A})$, with $b \geq a$, we always have $|x| \geq |y|$. Ordering based on ℓ_x and ℓ_y enables us to use different arguments in the two cases: when $\ell_y \leq \ell_x$, we count pairs based on x, and when $\ell_y > \ell_x$, we count pairs based on y.

3.3.1 The case $\ell_y \leq \ell_x$

Lemma 3.3.2. *Let b, a be nonnegative integers with $b \geq a$ and $1 \leq b + a \leq 2\log|\mathcal{A}|$.*

- *If $b + a$ is even, then*

$$|\{\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}) : \ \ell_y \leq \ell_x\}| \leq \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot (n \cdot \ell)^{(b+a)/2} \cdot |\mathcal{A}|.$$

- *If $b+a$ is odd, then*

$$|\{\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}) : \ \ell_y \leq \ell_x\}| \leq \left(\frac{4\sqrt{2}e}{b+a}\right)^{b+a} \cdot (n \cdot \ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}| \cdot |\mathcal{A}|.$$

Proof. Fix $x \in \mathcal{A}$. For each $p \in [a] \cup \{0\}$, we will bound the number of $y \in \{0,1\}^n$ such that $\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A})$ and $\ell_y \leq \ell_x$ and $|(y \setminus x) \cap \{\beta+1,\ldots,n\}| = p$. We claim that the number of such y is at most

$$\binom{n-\beta-\ell_x}{p}\binom{\ell_x}{p}\binom{\beta-|x|+\ell_x}{a-p}\binom{|x|}{b-p}. \tag{3.14}$$

Indeed, the first two factors count the ways to replace p elements in x with p new elements that are larger than β, and the final two factors count the ways to replace $b-p$ elements in x with $a-p$ new elements that are at most β.

Recall that Lemma 3.3.1 implies that $\ell_x \leq \ell$. Therefore, the quantity in (3.14) is at most

$$\binom{n}{p}\binom{\ell}{p}\binom{\beta}{a-p}\binom{|x|}{b-p} \ \leq \ \frac{(n\ell)^p \cdot \beta^{a-p}|x|^{b-p}}{(p!)^2 \cdot (a-p)! \cdot (b-p)!}. \tag{3.15}$$

We note that for $i,j \geq 0$ we have $i^i j^j \geq \left(\frac{i+j}{2}\right)^{i+j}$. Indeed, taking logs and dividing by 2, this is equivalent to

$$\tfrac{1}{2}(i\log i + j\log j) \geq \tfrac{i+j}{2}\log\left(\tfrac{i+j}{2}\right),$$

which follows from the convexity of the function $z \mapsto z\log z$. Hence, we may bound from

below the denominator of the right-hand side of equation (3.15) as follows:

$$(p!)^2 \cdot (a-p)! \cdot (b-p)! \geq \frac{p^{2p} \cdot (a-p)^{a-p} \cdot (b-p)^{b-p}}{e^{b+a}} \quad \text{(by Stirling's approximation)} \tag{3.16}$$

$$\geq \left(\frac{b+a}{4e}\right)^{b+a} \quad \text{(by two applications of } i^i j^j \geq \left(\frac{i+j}{2}\right)^{i+j}\text{)}. \tag{3.17}$$

We now break the bounding of (3.15) into two cases, based on the parity of $b+a$. For both cases, recall that Proposition 3.3.1 implies that $\beta|x| \leq n\ell$ and $\beta^2 \leq n\ell$ and $|x|^2 \leq n\ell$.

The case where $b+a$ is even. We bound the numerator of the RHS of (3.15) by

$$(n\ell)^p \cdot \beta^{a-p}|x|^{b-p} \leq (n\ell)^p \cdot (n\ell)^{(a-p)/2} \cdot (n\ell)^{(b-p)/2} = (n\ell)^{(b+a)/2}.$$

Summing the above bound on (3.15) over $p \in [a] \cup \{0\}$ and employing (3.17), we obtain

$$\begin{aligned} |\{y \in \mathcal{A} : \{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}),\ \ell_y \leq \ell_x\}| &\leq \sum_{p=0}^{a} \frac{(n\ell)^{(b+a)/2}}{(p!)^2 \cdot (b-p)! \cdot (a-p)!} \\ &\leq (a+1) \cdot \frac{(n\ell)^{(b+a)/2}\,(4e)^{(b+a)}}{(b+a)^{b+a}} \\ &\leq \frac{(n\ell)^{(b+a)/2}\,(4\sqrt{2}e)^{(b+a)}}{(b+a)^{b+a}}, \end{aligned}$$

where the last inequality uses the fact that $(a+1) \leq (\sqrt{2})^{b+a}$, leading to the factor $(4\sqrt{2}e)^{(b+a)}$.

The case where $b+a$ is odd. In this case, we have $b \geq a+1 \geq p+1$. We recall that $|x| \leq \log|\mathcal{A}|$, and we upper bound the numerator of the RHS of (3.15) by

$$(n\ell)^p \cdot \beta^{a-p}|x|^{b-p} \leq (n\ell)^p \cdot (n\ell)^{(a-p)/2} \cdot (n\ell)^{(b-p-1)/2} \cdot \log|\mathcal{A}| = (n\ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}|.$$

Summing the above bound on (3.15) over $p \in [a] \cup \{0\}$ and employing (3.17), we obtain

$$\begin{aligned} |\{y \in \mathcal{A} : \{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}),\ \ell_y \le \ell_x\}| &\le \sum_{p=0}^{a} \frac{(n\ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}|}{(p!)^2 \cdot (b-p)! \cdot (a-p)!} \\ &\le \frac{(n\ell)^{(b+a-1)/2} (4\sqrt{2}e)^{(b+a)} \cdot \log|\mathcal{A}|}{(b+a)^{b+a}}. \end{aligned}$$

In both even and odd cases, summing over $x \in \mathcal{A}$ completes the proof. □

3.3.2 The case $\ell_y > \ell_x$

Lemma 3.3.3. *Let b, a be nonnegative integers with $b \ge a$ and $1 \le b + a \le 2\log|\mathcal{A}|$.*

- *If $b + a$ is even, then*

$$|\{\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}) :\ \ell_y > \ell_x\}| \le \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot (n \cdot \ell)^{(b+a-2)/2} \cdot \ell\beta \cdot |\mathcal{A}|.$$

- *If $b + a$ is odd, then*

$$|\{\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}) :\ \ell_y > \ell_x\}| \le \left(\frac{4\sqrt{2}e}{b+a}\right)^{b+a} \cdot (n \cdot \ell)^{(b+a-1)/2} \cdot \ell \cdot |\mathcal{A}|.$$

Proof. Fix $y \in \mathcal{A}$. For each $p \in [a]$, we will bound the number of $x \in \{0,1\}^n$ such that $\{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A})$ and $\ell_y > \ell_x$ and $|(x \setminus y) \cap \{\beta+1, \ldots, n\}| = p - 1$. We claim that the number of such x is at most

$$\binom{n-\beta-\ell_y}{p-1}\binom{\ell_y}{p}\binom{\beta-|x|+\ell_y}{b-p+1}\binom{|y|}{a-p}. \tag{3.18}$$

Indeed, the first two factors count the ways to replace p elements in y with $p-1$ new elements that are larger than β, and the final two factors count the ways to replace $a-p$ elements in y with $b-p+1$ new elements that are at most β.

Recall that Lemma 3.3.1 implies that $\ell_y \leq \ell$. Thus, the quantity in (3.18) is at most

$$\binom{n}{p-1}\binom{\ell}{p}\binom{\beta}{b-p+1}\binom{|y|}{a-p} \;\leq\; \frac{(n\ell)^{p-1}\cdot \ell \cdot \beta^{b-p+1}\cdot |y|^{a-p}}{(p-1)!\cdot p!\cdot (b-p+1)!\cdot (a-p)!}. \tag{3.19}$$

Similarly to in the proof of Lemma 3.3.2 (i.e., by applying Stirling's approximation and the fact $i^i j^j \geq (\frac{i+j}{2})^{i+j}$), we lower bound the denominator of the right hand side of (3.19) as follows:

$$\begin{aligned}(p-1)!\cdot p!\cdot (b-p+1)!\cdot (a-p)! &\geq \frac{(p-1)^{p-1}\cdot p^p\cdot (a-p)^{a-p}\cdot (b-p+1)^{b-p+1}}{e^{b+a}}\\ &\geq \left(\frac{b+a}{4e}\right)^{b+a}. \end{aligned}\tag{3.20}$$

Recall that Proposition 3.3.1 implies that $\beta^2 \leq n\ell$ and $|y|^2 \leq n\ell$. We now break into two cases, based on the parity of $b+a$.

The case where $b+a$ is even. Notice that $\ell_y > \ell_x$ and $|x| \geq |y|$ implies $a \geq 1$ and $b+a \geq 2$. We upper bound the numerator of the RHS of (3.19) by

$$(n\ell)^{p-1}\cdot \ell \cdot \beta^{b-p+1}\cdot |y|^{a-p} \leq (n\ell)^{p-1}\cdot \ell \cdot \beta \cdot (n\ell)^{(b-p)/2}\cdot (n\ell)^{(a-p)/2} = (n\ell)^{(b+a-2)/2}\cdot \ell\beta.$$

Summing our bound on (3.19) over $p \in [a]$, employing (3.20), and using that $a \leq (\sqrt{2})^{b+a}$,

$$\begin{aligned}|\{x \in \mathcal{A} : \{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}),\ \ell_y > \ell_x\}| \;&\leq\; \sum_{p=1}^{a} \frac{(n\ell)^{(b+a-2)/2}\cdot \ell\beta}{p!\cdot (p-1)!\cdot (b-p+1)!\cdot (a-p)!}\\ &\leq\; \frac{(n\ell)^{(b+a-2)/2}\,(4\sqrt{2}e)^{(b+a)}\cdot \beta\ell}{(b+a)^{b+a}}.\end{aligned}$$

The case where $b+a$ is odd. Notice that $\ell_y > \ell_x$ and $|x| \geq |y|$ implies $a \geq 1$, and in this case, $b \geq a+1 \geq p+1$. We upper bound the RHS of (3.19) by

$$(n\ell)^{p-1}\cdot \ell \cdot \beta^{b-p+1}\cdot |y|^{a-p} \leq (n\ell)^{p-1}\cdot \ell \cdot (n\ell)^{(b-p+1)/2}\cdot (n\ell)^{(a-p)/2} = (n\ell)^{(b+a-1)/2}\cdot \ell.$$

Summing our bound on (3.19) over $p \in [a]$, employing (3.20), and using that $a \leq (\sqrt{2})^{b+a}$,

$$\begin{aligned} |\{x \in \mathcal{A} : \{x,y\} \in \mathsf{E}_{(b,a)}(\mathcal{A}),\ \ell_y > \ell_x\}| &\leq \sum_{p=1}^{a} \frac{(n\ell)^{(b+a-1)/2} \cdot \ell}{p! \cdot (p-1)! \cdot (b-p+1)! \cdot (a-p)!} \\ &\leq \frac{(n\ell)^{(b+a-1)/2} (4\sqrt{2}e)^{(b+a)} \cdot \ell}{(b+a)^{b+a}}. \end{aligned}$$

In both even and odd cases, summing over $y \in \mathcal{A}$ completes the proof. □

3.3.3 Finishing the proof

Proof of Theorem 3.1.5. Recall that $\mathcal{U} := \{(b,a) \in \mathbb{Z}^2_{\geq 0} :\ b \geq a \text{ and } b+a \leq 2t\}$. Invoking (3.11) and using Lemma 3.3.2 and Lemma 3.3.3, we will upper bound each term in

$$\mathsf{e}_{\leq 2t}(\mathcal{A}) = \sum_{(b,a)\in\mathcal{U}} \mathsf{e}_{(b,a)}(\mathcal{A}).$$

For all $(b,a) \in \mathcal{U}$, we claim that

$$\frac{\mathsf{e}_{(b,a)}(\mathcal{A})}{|\mathcal{A}|} \leq \left(\frac{4e}{t}\right)^{2t} (n\ell)^t. \tag{3.21}$$

Assuming that (3.21) holds, and using that $|\mathcal{U}| \leq 2^{2t}$, we have

$$\sum_{(b,a)\in\mathcal{U}} \frac{\mathsf{e}_{(b,a)}(\mathcal{A})}{|\mathcal{A}|} \leq |\mathcal{U}| \cdot \left(\frac{4e}{t}\right)^{2t} (n\ell)^t \leq \left(\frac{8e}{t}\right)^{2t} (n\ell)^t,$$

which implies the bound in the theorem statement. To prove (3.21), we will use Proposition 3.3.1 and the fact that $t \leq \lfloor \log |\mathcal{A}| \rfloor$. When $b+a$ is even, then combining Lemma 3.3.2 and Lemma 3.3.3 (using $\beta\ell \leq n\ell$), we have

$$\begin{aligned}
e_{(b,a)}(\mathcal{A}) &\leq \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot (n\ell)^{(b+a)/2} \cdot |\mathcal{A}| + \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot (n\ell)^{(b+a-2)/2} \cdot \ell\beta \cdot |\mathcal{A}| \\
&= \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a-2)/2} \cdot (n\ell + \ell\beta) \\
&\leq 2 \cdot \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a)/2} \quad (\text{as } \ell\beta \leq n\ell) \\
&\leq \left(\frac{8e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a)/2} \quad (\text{as } 2 \leq \sqrt{2}^{(b+a)}).
\end{aligned}$$

To verify (3.21), it suffices to show that the RHS of the above inequality increases with $b+a$ (i.e., that it is maximized over $\mathcal{U}$ at $b+a=2t$). Indeed, let $k = b+a \geq 2$. Then, it suffices to show that

$$\left(\frac{8e}{k-1}\right)^{k-1} \cdot (n \cdot \ell)^{k/2-1/2} \leq \left(\frac{8e}{k}\right)^{k} \cdot (n \cdot \ell)^{k/2}. \tag{3.22}$$

After rearranging, we have

$$\frac{k}{8e}\left(\frac{k}{k-1}\right)^{k-1} \;\leq\; \frac{k}{8} \;\leq\; (n\ell)^{1/2},$$

where the first inequality uses that $(\frac{k}{k-1})^{k-1} \leq e$, and the second inequality uses that $(k/8)^2 \leq t^2 \leq \lfloor \log |\mathcal{A}| \rfloor^2 \leq n\ell$, which holds by Proposition 3.3.1 (v).

Similarly, when $b+a$ is odd, Lemma 3.3.2 and Lemma 3.3.3 (using $\ell \leq \log |\mathcal{A}|$) imply

that

$$\begin{aligned}
e_{(b,a)}(\mathcal{A}) &\leq \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} (n\ell)^{(b+a-1)/2} \log|\mathcal{A}| \cdot |\mathcal{A}| + \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot (n\ell)^{(b+a-1)/2} \ell |\mathcal{A}| \\
&= \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a-1)/2} \cdot (\log|\mathcal{A}| + \ell) \\
&\leq 2 \cdot \left(\frac{4\sqrt{2}e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}| \quad (\text{as } \ell \leq \log|\mathcal{A}|) \\
&\leq \left(\frac{8e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}| \quad (\text{as } 2 \leq \sqrt{2}^{(b+a)}).
\end{aligned}$$

We claim that $\left(\frac{8e}{b+a}\right)^{(b+a)} \cdot |\mathcal{A}| \cdot (n\ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}|$ is maximised over $\mathcal{U}$ when $b+a = 2t-1$. Indeed, letting $k = b+a \geq 2$, we have

$$\begin{aligned}
&\left(\frac{8e}{k-1}\right)^{k-1} \cdot |\mathcal{A}| \cdot (n\ell)^{(k-2)/2} \cdot \log|\mathcal{A}| \leq \left(\frac{8e}{k}\right)^{k} \cdot |\mathcal{A}| \cdot (n\ell)^{(k-1)/2} \cdot \log|\mathcal{A}| \\
\iff &\left(\frac{k}{k-1}\right)^{k-1} \frac{k}{8e} \leq (n\ell)^{1/2},
\end{aligned}$$

where the last inequality holds since $(k/8)^2 \leq t^2 \leq \lfloor \log|\mathcal{A}| \rfloor^2 \leq n\ell$, by Proposition 3.3.1 (v) and $\left(\frac{k}{k-1}\right)^{k-1} \leq e$. It follows that

$$\begin{aligned}
e_{(b,a)}(\mathcal{A}) &\leq \left(\frac{8e}{2t-1}\right)^{(2t-1)} \cdot |\mathcal{A}| \cdot (n\ell)^{t-1} \cdot \log|\mathcal{A}| \\
&= \left(\frac{4e}{t}\right)^{2t} \cdot |\mathcal{A}| \cdot (n\ell)^{t} \cdot \log|\mathcal{A}| \cdot \left(\frac{2t}{2t-1}\right)^{(2t-1)} \cdot \frac{t}{4e} \cdot \frac{1}{n\ell} \\
&\leq \left(\frac{4e}{t}\right)^{2t} \cdot |\mathcal{A}| \cdot (n\ell)^{t} \cdot \log|\mathcal{A}| \cdot \frac{t}{4} \cdot \frac{1}{n\ell} \\
&\leq \left(\frac{4e}{t}\right)^{2t} \cdot |\mathcal{A}| \cdot (n\ell)^{t},
\end{aligned}$$

where the last inequality follows from noting that $\frac{t \log|\mathcal{A}|}{4} \leq \lfloor \log|\mathcal{A}| \rfloor \log|\mathcal{A}| \leq n\ell$ (which follows from Proposition 3.3.1 (v)).

□

3.4 The general case for odd distances

Proof of Theorem 3.1.6. The following proof has very similar structure to the proof of Theorem 3.1.5, so we omit detailed calculations.

Using the notation defined above, it is required to prove that if $\mathcal{A} \subset \{0,1\}^n$ and $t \in \mathbb{N}$ with $t \leq \log|\mathcal{A}|$, then

$$|\mathsf{E}_{\leq 2t+1}(\mathcal{A})| := \mathsf{e}_{\leq 2t+1}(\mathcal{A}) \;\leq\; \left(\frac{16e}{2t+1}\right)^{2t+1} \cdot (n\cdot\ell)^t \cdot |\mathcal{A}| \cdot \log|\mathcal{A}|.$$

Letting $\mathcal{U}' = \{(b,a) \in \mathbb{Z}^2_{\geq 0} :\ b \geq a \text{ and } b+a \leq 2t+1\}$, observe that

$$\mathsf{e}_{\leq 2t+1}(\mathcal{A}) = \sum_{(b,a)\in\mathcal{U}'} \mathsf{e}_{(b,a)}(\mathcal{A}).$$

We will upper bound each term in the above sum. For $(b,a) \in \mathcal{U}'$, we claim that

$$\frac{\mathsf{e}_{(b,a)}(\mathcal{A})}{|\mathcal{A}|} \leq 2\left(\frac{4\sqrt{2}e}{2t+1}\right)^{2t+1} (n\ell)^t \cdot \log|\mathcal{A}| \leq \left(\frac{8e}{2t+1}\right)^{2t+1} (n\ell)^t \cdot \log|\mathcal{A}|. \tag{3.23}$$

Assuming that (3.23) holds, and using that $|\mathcal{U}'| \leq 2^{2t+1}$, we have

$$\sum_{(b,a)\in\mathcal{U}} \frac{\mathsf{e}_{(b,a)}(\mathcal{A})}{|\mathcal{A}|} \leq |\mathcal{U}'| \cdot \left(\frac{8e}{2t+1}\right)^{2t+1} (n\ell)^t \cdot \log|\mathcal{A}| \;\leq\; \left(\frac{16e}{2t+1}\right)^{2t+1} (n\ell)^t \cdot \log|\mathcal{A}|,$$

which establishes the bound in the theorem statement.

We now prove (3.23). When $b+a$ is even, then $b+a \leq 2t$ and (3.23) follows from (3.21). When $b+a$ is odd, then Lemma 3.3.2 and Lemma 3.3.3 (using $\ell \leq \log|\mathcal{A}|$) imply that

$$\frac{\mathsf{e}_{(b,a)}(\mathcal{A})}{|\mathcal{A}|} \leq \left(\frac{8e}{b+a}\right)^{b+a} \cdot (n\cdot\ell)^{(b+a-1)/2} \cdot \log|\mathcal{A}| \leq \left(\frac{8e}{2t+1}\right)^{2t+1} \cdot (n\cdot\ell)^t \cdot \log|\mathcal{A}|,$$

where we use that the quantity $\left(\frac{8e}{b+a}\right)^{b+a} \cdot (n\cdot\ell)^{(b+a-1)/2}$ increases with $b+a$ (and is maximized over $\mathcal{U}'$ at $b+a = 2t+1$), analogous to the proof of (3.22). $\square$

3.5 Technical results

Here we provide proof of the technical proposition, Proposition 3.1.3. For this we need the following tool.

Proposition 3.5.1. *Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be defined as follows*

$$f(x) = \begin{cases} \left(\frac{x}{m}\right)^m + \left(\frac{x}{m+1}\right)^{m+1} - e^{x/e} & \text{if } x \in [me, (m+1)e), \text{ for some } m \in \mathbb{N}, m \geq 1 \\ 1 + x - e^{x/e} & \text{if } x \in [0, e) \end{cases}$$

Then the following hold.

(1) For $x \in [0, e)$, $f(x) \geq x/e \geq 0$.

(2) For $x \in [e, 2e)$, $f(x) \geq \frac{e^2}{4} + (2 - \frac{e}{4})(x - e) \geq 0$.

(3) For $m \geq 2$ and $x \in [me, (m+1)e)$, we have

$$e^{x/e} - \min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} \leq \frac{1}{m}\min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\},$$

from which it immediately follows that

$$f(x) \geq \max\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} - \frac{1}{m}\min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} \geq 0.$$

Proof. We split our proof into parts for each of the statements.

Part (1). Suppose first that $x \in [0, e)$, so $f(x) = 1 + x - e^{x/e}$. Then $\frac{d^2 f}{dx^2} = -e^{x/e-2} < 0$ and so f is concave in this range. Hence, we have

$$f(x) \geq f(0) + \frac{f(e) - f(0)}{e - 0}x = \frac{x}{e},$$

as required.

Part (2). Suppose next that $x \in [e, 2e)$, so that $f(x) = x + \frac{x^2}{4} - e^{x/e}$. We let

$$g(x) = f(x) - (\frac{e^2}{4} + (2 - \frac{e}{4})(x - e)) = (2e - \frac{e^2}{2}) + (-1 + \frac{e}{4})x + \frac{x^2}{4} - e^{x/e},$$

and note the following:

$$\begin{aligned} g'(x) &= (-1 + \frac{e}{4}) + \frac{x}{2} - e^{x/e-1} \\ g''(x) &= \frac{1}{2} - e^{x/e-2} \\ g(e) &= g(2e) = 0. \end{aligned}$$

Clearly, $g''(x)$ is decreasing in x and has a unique root at $x = e(2 - \ln(2))$. Therefore $g''(x) > 0$ for $x \in [e, e(2 - \ln(2)))$ and $g''(x) < 0$ for $x \in (e(2 - \ln(2)), 2e)$. We also note that $g'(e) = \frac{3e}{4} - 2 > 0$, $g'(e(2 - \ln(2))) = -1 + \frac{3-2\ln(2)}{4}e > 0$ and $g'(2e) = -1 + \frac{e}{4} < 0$.

As $g''(x) < 0$ for $x \in (e(2-\ln(2)), 2e)$ and $g'(e(2-\ln(2)))g'(2e) < 0$ we see that $g'(x) = 0$ has a unique root in $(e(2 - \ln(2)), 2e)$. In addition, $g''(x) > 0$ for $x \in [e, e(2 - \ln(2)))$ and $g'(e)g'(e(2 - \ln(2))) > 0$ so we see that $g'(x) = 0$ has no solutions in $[e, e(2 - \ln(2))]$. Hence $g(x)$ has a unique maximum in $[e, 2e)$, and no other stationary points. From this, and the fact that $g(e) = g(2e) = 0$ we deduce that $g(x) \geq 0$ for all $x \in [e, 2e)$. This shows that

$$f(x) \geq \frac{e^2}{4} + (2 - \frac{e}{4})(x - e)$$

for $x \in [e, 2e)$, as claimed.

Part (3). Suppose finally that $x \in [me, (m + 1)e)$ for some $2 \leq m \in \mathbb{N}$. We now split into two cases: the case $\left(\frac{x}{m}\right)^m \geq \left(\frac{x}{m+1}\right)^{m+1}$, and the case $\left(\frac{x}{m}\right)^m < \left(\frac{x}{m+1}\right)^{m+1}$.

Case 1: Suppose first that the former case holds. Then

$$\begin{aligned} e^{x/e} - \min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} &= e^{x/e} - \left(\frac{x}{m+1}\right)^{m+1} \\ &= -\int_{t=x/e}^{m+1} \left(\frac{x}{t}\right)^t (\ln\left(\frac{x}{t}\right) - 1)dt \\ &= \int_{t=x/e}^{m+1} \left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right) dt \\ &\leq (m+1-x/e) \max_{t\in[x/e,m+1]} \left\{\left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right)\right\}. \end{aligned}$$

To bound $\max_{t\in[x/e,m+1]} \left\{\left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right)\right\}$ we show the maximum is attained at $t = m+1$. Indeed, differentiating with respect to t we get:

$$\begin{aligned} \frac{d}{dt}\left(\left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right)\right) &= \left(\frac{x}{t}\right)^t \left(\frac{1}{t} - \ln\left(\frac{t}{x/e}\right)^2\right) \\ &\geq \left(\frac{x}{t}\right)^t \left(\frac{1}{m+1} - \left(\ln\left(\frac{m+1}{x/e}\right)\right)^2\right). \end{aligned}$$

It is a standard fact that for $y > 0$ we have $\frac{y-1}{y} \leq \ln(y) \leq y - 1$. Noting that $\frac{m+1}{x/e} > 0$, we apply this fact to see:

$$\ln\left(\frac{m+1}{x/e}\right) \leq \frac{m+1}{x/e} - 1 = \frac{(m+1) - x/e}{x/e} \leq e/x.$$

Hence, we have

$$\begin{aligned} \frac{d}{dt}\left(\left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right)\right) &\geq \left(\frac{x}{t}\right)^t \left(\frac{1}{m+1} - (e/x)^2\right) \\ &= \left(\frac{x}{t}\right)^t \left(\frac{(x/e)^2 - (m+1)}{(m+1)(x/e)^2}\right) \\ &\geq \left(\frac{x}{t}\right)^t \left(\frac{m^2 - m - 1}{(m+1)(x/e)^2}\right) \geq 0, \end{aligned}$$

where the final inequality holds since $m \geq 2$. Thus $\left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right)$ is increasing on the interval $t \in [x/e, m+1]$, and attains its maximum at $t = m+1$. Therefore, we may

bound the integral as follows:

$$\int_{t=x/e}^{m+1} \left(\frac{x}{t}\right)^t \ln\left(\frac{t}{x/e}\right) dt \leq (m+1-x/e)\left(\frac{x}{m+1}\right)^{m+1} \ln\left(\frac{m+1}{x/e}\right) \leq \left(\frac{x}{m+1}\right)^{m+1} \frac{1}{m}.$$

The final inequality holds as $(m+1-x/e) \leq 1$ and $\ln\left(\frac{m+1}{x/e}\right) \leq \frac{1}{m}$. The first of these is trivial, and the second can be seen as follows. We define $\varepsilon \in [0,1)$ by $x = (m+\varepsilon)e$, then

$$\ln\left(\frac{m+1}{x/e}\right) = \ln\left(\frac{m+1}{m+\varepsilon}\right) \leq \frac{1-\varepsilon}{m+\varepsilon} \leq \frac{1}{m}.$$

Hence, we have shown that

$$e^{x/e} - \min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} \leq \left(\frac{x}{m+1}\right)^{m+1} \frac{1}{m},$$

i.e., that the claim holds in the former case.

Case 2: Suppose secondly that the latter case holds. Then we have

$$\begin{aligned} e^{x/e} - \min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} &= e^{x/e} - \left(\frac{x}{m}\right)^m \\ &= \int_{t=m}^{x/e} \left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right) dt \\ &\leq (x/e-m) \max_{t\in[m,x/e]} \left\{\left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right)\right\}. \end{aligned}$$

To bound $\max_{t\in[m,x/e]} \left\{\left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right)\right\}$ we show that the maximum is attained at $t = m$. Differentiating with respect to t we get:

$$\begin{aligned} \frac{d}{dt}\left(\left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right)\right) &= \left(\frac{x}{t}\right)^t \left(\ln\left(\frac{x/e}{t}\right)^2 - \frac{1}{t}\right) \\ &\leq \left(\frac{x}{t}\right)^t \left(\left(\ln\left(\frac{x/e}{m}\right)\right)^2 - \frac{1}{x/e}\right). \end{aligned}$$

Observe that

$$\ln\left(\frac{x/e}{m}\right) \leq \frac{x/e}{m} - 1 = \frac{(x/e)-m}{m} \leq \frac{1}{m}.$$

Substituting this bound into the previous equation gives

$$\begin{aligned}\frac{d}{dt}\left(\left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right)\right) &\leq \left(\frac{x}{t}\right)^t \left(\left(\frac{1}{m}\right)^2 - \frac{1}{x/e}\right) \\ &= \left(\frac{x}{t}\right)^t \left(\frac{x/e - m^2}{m^2(x/e)}\right) \\ &\leq \left(\frac{x}{t}\right)^t \left(\frac{m+1-m^2}{m^2(x/e)}\right) \leq 0.\end{aligned}$$

(Note that the final inequality holds as $m \geq 2$.) Hence, $\left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right)$ is non-increasing on the interval $t \in [m, x/e]$, and so attains its maximum at $t = m$. We may bound the integral as follows:

$$\int_{t=m}^{x/e} \left(\frac{x}{t}\right)^t \ln\left(\frac{x/e}{t}\right) dt \leq (x/e - m)\left(\frac{x}{m}\right)^m \ln\left(\frac{x/e}{m}\right) \leq \left(\frac{x}{m}\right)^m \frac{1}{m}.$$

(Note that the final inequality holds as $((x/e) - m) \leq 1$ and $\ln\left(\frac{x/e}{m}\right) \leq \frac{1}{m}$. The first of these is trivial, and the second can be seen as follows. We define $\varepsilon \in [0, 1)$ by $x = (m+\varepsilon)e$. Then

$$\ln\left(\frac{x/e}{m}\right) = \ln\left(\frac{m+\varepsilon}{m}\right) \leq \frac{\varepsilon}{m} \leq \frac{1}{m}.$$

Hence, we have shown that

$$e^{x/e} - \min\left\{\left(\frac{x}{m}\right)^m, \left(\frac{x}{m+1}\right)^{m+1}\right\} \leq \left(\frac{x}{m}\right)^m \frac{1}{m},$$

i.e., that the claim holds in the latter case. This completes the proof of the claim. □

We now prove Proposition 3.1.3.

Proof of Proposition 3.1.3. Fix $m \in \mathbb{N}, K \in \mathbb{R}^+$ and consider $\left(\frac{K}{m+\lambda}\right)^{m+\lambda}$. Differentiating this with respect to λ we find:

$$\frac{d}{d\lambda}\left(\left(\frac{K}{m+\lambda}\right)^{m+\lambda}\right) = \left(\frac{K}{m+\lambda}\right)^{m+\lambda}\left(\ln\left(\frac{K/e}{m+\lambda}\right)\right).$$

The only solution to $\frac{d}{d\lambda}\left(\left(\frac{K}{m+\lambda}\right)^{m+\lambda}\right) = 0$ is $\lambda = \frac{K}{e} - m$.

If $\frac{K}{e} - m < 0$, then for all $\lambda \in [0,1)$ we have $\frac{K/e}{m+\lambda} < \frac{m}{m+\lambda} \leq 1$, so the derivative is negative, and the maximum is attained by $\left(\frac{K}{m}\right)^m$, so the claim holds in this case.

If $\frac{K}{e} - m \geq 1$, then for all $\lambda \in [0,1)$ we have $\frac{K/e}{m+\lambda} \geq \frac{m+1}{m+\lambda} > 1$, so the derivative is positive, and the maximum is attained by $\left(\frac{K}{m+1}\right)^{m+1}$, so the claim holds in this case also.

Finally, suppose that $\frac{K}{e} - m \in [0,1)$. Then the maximum is at $\lambda = \frac{K}{e} - m$, but we appeal to Proposition 3.5.1 to get

$$\left(\frac{K}{m}\right)^m + \left(\frac{K}{m+1}\right)^{m+1} - \left(\frac{K}{m+\lambda}\right)^{m+\lambda} = \left(\frac{K}{m}\right)^m + \left(\frac{K}{m+1}\right)^{m+1} - e^{K/e} = f(K) \geq 0.$$

This leaves the case $m = 0$, which we resolve similarly. First, we differentiate $(K/\lambda)^\lambda$ with respect to λ to get

$$\frac{d}{d\lambda}\left(\left(\frac{K}{\lambda}\right)^\lambda\right) = \left(\frac{K}{\lambda}\right)^\lambda \left(\ln\left(\frac{K/e}{\lambda}\right)\right),$$

and note that

(1) the derivative has a unique root at $\lambda = K/e$,

(2) the derivative is strictly positive if $\lambda < K/e$,

(3) the derivative is strictly negative if $\lambda > K/e$.

Consequently, if $K/e \geq 1$, then $\left(\frac{K}{\lambda}\right)^\lambda \leq K$ for all $\lambda \in [0,1)$, so the claim holds. If $0 < K/e < 1$ then $\left(\frac{K}{\lambda}\right)^\lambda \leq e^{K/e}$, so by Proposition 3.5.1

$$1 + K - \left(\frac{K}{\lambda}\right)^\lambda \geq 1 + K - e^{K/e} = f(K) \geq 0.$$

This completes the proof. □

3.6 Some open questions

One obvious open problem is to prove exact edge isoperimetric inequalities for the graphs we consider. It would also be interesting to study graphs on $[k]^n$ induced by various natural metrics, for $k \geq 3$. Two possible generalizations of our results would be for the families of graphs connecting pairs in $[k]^n$ either with ℓ_1-distance at most r, or Hamming distance at most r. Bollobás and Leader [16] and Clements and Lindström [20] have solved the respective distance one cases.

3.7 Acknowledgments

The work in this chapter is based on the preprint "Edge Isoperimetric Inequalities for Powers of the Hypercube", written jointly by Cyrus Rashtchian and the author.

Chapter 4

Families of sets that are pairwise close

4.1 Introduction

For a positive integer n, we define the *cyclic distance* on $[n] := \{1, 2, ..., n\}$ by

$$\text{dist} : [n] \times [n] \to \mathbb{Z}_{\geq 0}; \quad \text{dist}(a, b) = \min_{z \in \mathbb{Z}} \{|z| : b \equiv a + z \mod n\},$$

that is the shortest cyclic distance from a to b when $[n]$ is used to label n regularly spaced points on a circle. Given a set X we let $\mathcal{P}(X) := \{A \subseteq X\}$, i.e., the power set of X. We will often identify $\mathcal{P}([n])$ with $\{0, 1\}^n$ via the bijection $A \leftrightarrow (x_i)_{i=1}^n$ such that $x_i = 1$ if and only if $i \in A$. For a non-negative integer $k \leq n$ we let $[n]^{(k)} := \{A \subseteq [n] : |A| = k\}$.

Let d be a non-negative integer. We say a pair of subsets $A, B \in \mathcal{P}([n])$ is *d-close* if $\min_{a \in A, b \in B} \text{dist}(a, b) \leq d$. Furthermore, we say a set system $\mathcal{A} \subseteq \mathcal{P}([n])$ is d-close if every pair of sets $A, B \in \mathcal{A}$ is d-close. We also say that a pair of set systems $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}([n])$ is *cross d-close* if for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$ the pair A, B is d-close. The main purpose of this chapter is to prove upper bounds on the maximum possible size of d-close set

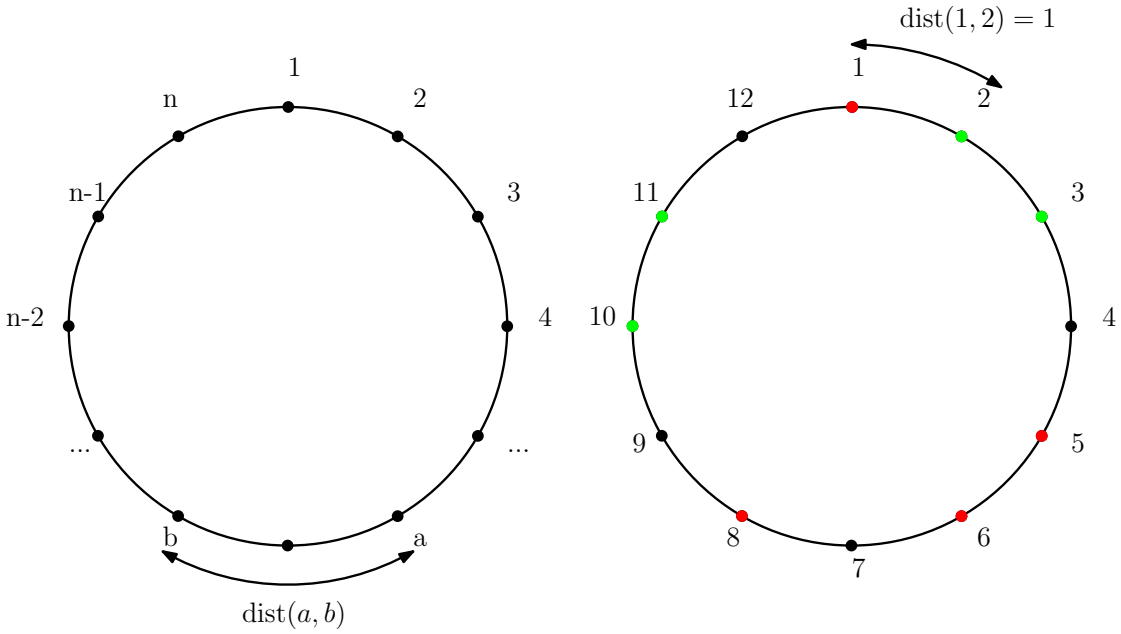


Figure 4.1: (*left*) The cyclic distance between $a, b \in [n]$ is the number of steps in the shortest arc between a and b when considering $[n]$ as ordered, equally spaced points around a circle. (*right*) An example of a 1-close pair of subsets of $[12]$, the red set $\{1, 5, 6, 8\}$ and the green set $\{2, 3, 10, 11\}$ are 1-close as $\text{dist}(1, 2) = 1$.

systems and cross d-close pairs of set systems.

Intersection problems are an extensively studied class of problems from Extremal Combinatorics, asking for the maximal possible size of a family of combinatorial objects subject to conditions on the pairwise intersections of objects in the family. This area of study was introduced in 1961 by Erdős, Ko and Rado [28], when they proved that for every $k \leq n/2$ an *intersecting* family $\mathcal{F} \subseteq [n]^{(k)}$ (i.e., $\mathcal{F} \subseteq [n]^{(k)}$ such that every pair $A, B \in \mathcal{F}$ has $A \cap B \neq \emptyset$) can have size at most $\binom{n-1}{k-1}$, and furthermore that if $k < n/2$ and $|\mathcal{F}|$ is maximum then $\mathcal{F}$ is isomorphic to $\{A \in [n]^{(k)} \ : \ 1 \in A\}$.

Since the original Erdős-Ko-Rado theorem, many intersection theorems have been proved. Some examples are:

- The Ahlswede-Khachatrian theorem [6, 7] finding for all positive integers $1 \leq t \leq k \leq n$ the exact maximum possible size of families $\mathcal{F} \subseteq [n]^{(k)}$ in which any pair of sets $A, B \in \mathcal{F}$ have $|A \cap B| \geq t$. This result is a remarkably precise extension of an

original theorem of Erdős, Ko and Rado which showed that for $k \in \mathbb{N}$ and $t \in [k]$ there exists $n_0(k,t)$ such that if $n \geq n_0(k,t)$ then a t-intersecting k-uniform family of subsets of $[n]$ has size at most $\binom{n-t}{k-t}$.

- The result of Talbot [71] finding the maximum possible size of an intersecting uniform family of sets that are *separated*, where $A \subseteq [n]$ is 1-separated if for every pair $a, b \in A$ are separated by a cyclic gap of at least 1. For $n \in \mathbb{N}$ and $r \leq n/2$ the maximum size is that of the family $\{A \in [n]^{(r)} : A \text{ is separated}, 1 \in A\}$ which is the unique extremal family up to isomorphism.

- The structural result of Hilton and Milner [44] which determined the largest intersecting families $\mathcal{F} \subseteq [n]^{(k)}$ that are not isomorphic to $\{A \in [n]^{(k)} : 1 \in A\}$. In particular, Hilton and Milner proved that if $4 \leq 2k \leq n$ are positive integers, and $\mathcal{F} \subseteq [n]^{(k)}$ is an intersecting family such that $\bigcap_{A \in \mathcal{F}} A = \emptyset$, then $|\mathcal{F}| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$. This upper bound is witnessed by the family $\{A \in [n]^{(k)} : A = [k] \text{ or } n \in A, A \cap [k] \neq \emptyset\}$.

- The theorem of Ellis, Filmus and Friedgut [25] on triangle-intersecting families of graphs. For $n \in \mathbb{N}$, a family of graphs $\mathcal{G}$ on vertex set $[n]$ is said to be triangle intersecting if for each pair $G, H \in \mathcal{G}$, the intersection $G \cap H$ (i.e., the graph of edges common to G and H) contains a triangle. Ellis, Filmus and Friedgut proved that a triangle-intersecting family of graphs on vertex set $[n]$ can have size at most $2^{\binom{n}{2}-3}$, and this maximum is witnessed by $\{G \subseteq [n]^{(2)} : \{1,2\}, \{1,3\}, \{2,3\} \in G\}$, i.e., the collection of all graphs which contain some fixed triangle.

- A result of Ellis, Friedgut and Pilpel [26] which builds on work of Cameron and Yu [18] and Deza and Frankl [22] on intersecting families of permutations. For $n, k \in \mathbb{N}$, let S_n be the group of permutations of $[n]$, and say that a collection $I \subseteq S_n$ is k-intersecting if for all $\sigma, \pi \in I$ there exist distinct $i_1, ..., i_k \in [n]$ such that $\sigma(i_t) = \pi(i_t)$ for $t = 1, 2, ..., k$. Ellis, Friedgut and Pilpel proved a conjecture of Deza and Frankl that for each $k \in \mathbb{N}$, if n is sufficiently large with respect to k

then the largest possible k-intersecting families of permutations in S_n are cosets of stabilisers of k points.

- A number of results of Chung, Graham, Frankl and Shearer [19], including their result for set systems $\mathcal{F} \subseteq \mathcal{P}([n])$ such that for each $A, B \in \mathcal{F}$ the intersection $A \cap B$ contains a cyclic translate of the block $\{1, 2, ..., t\}$. They proved that such a family has size at most 2^{n-t}, witnessed by the family $\{A \subset [n] \ : \ \{1, 2, ..., t\} \subseteq A\}$.

The problem investigated in this chapter is another natural generalisation of the Erdős-Ko-Rado theorem [28] for *intersecting families*, i.e., set systems $\mathcal{A}$ such that for every pair of sets $A, B \in \mathcal{A}$ the intersection $A \cap B$ is nonempty. In particular they proved the following tight upper bounds:

- for $\mathcal{A} \subseteq \mathcal{P}([n])$ intersecting, $|\mathcal{A}| \leq 2^{n-1}$,
- for $\mathcal{A} \subseteq [n]^{(k)}$ intersecting, $|\mathcal{A}| \leq \binom{n-1}{k-1}$.

We note that the d-close definitions generalise the notions of intersecting families. Indeed, in the case $d = 0$ a pair of sets $A, B \subseteq [n]$ is 0-close if and only if $A \cap B \neq \emptyset$, i.e., 0-close set systems are precisely intersecting families. Our goal then is to generalise the above upper bounds. The main theorem of the chapter is the generalisation of the second inequality.

Remark 4.1.1. *We note that although 0-close set systems in $\mathcal{P}([n])$ have size at most $2^{n-1} = \frac{1}{2}|\mathcal{P}([n])|$, even 1-close set systems in $\mathcal{P}([n])$ can be very large, having size up to $(1 - o_n(1))2^n$. Indeed, consider the following construction due to Leader (private communication):*

$$\mathcal{A} = \left\{ A \subseteq [n] \mid A \cap \{2i-1, 2i\} \neq \emptyset \text{ for strictly more than half the } i \in \left\{1, 2, ..., \left\lfloor \frac{n}{2} \right\rfloor\right\}\right\}.$$

Then $\mathcal{A}$ is indeed 1-close, and by standard Chernoff bound inequalities:

$$|\mathcal{A}| \geq (1 - o_n(1))2^n.$$

Since it is clear that $|\mathcal{A}| \leq 2^n$, the above remark shows that this upper bound is essentially tight. Thus for the rest of the chapter we restrict our attention to uniform d-close set systems, i.e., we let $k \leq n$ be a non-negative integer and consider set systems $\mathcal{A} \subseteq [n]^{(k)}$ that are d-close, or pairs of set systems $\mathcal{A}, \mathcal{B} \subseteq [n]^{(k)}$ that are cross d-close.

We can also see that a slight variation of this construction works for d-close, k-uniform set systems when d is fixed and k is sufficiently large with respect to n in a way we will now make clear. Indeed, we fix $d \in \mathbb{N}$ and take $n \in \mathbb{N}$. Let $n = m(d+1)+r$ where $m, r \in \mathbb{N}$ such that $0 \leq r < d+1$. Now let $I_j = \{(j-1)(d+1)+1, (j-1)(d+1)+2, ..., j(d+1)\}$ for $j = 1, 2, ..., m$ and let

$$\mathcal{A} = \{A \in [n]^{(k)} \mid A \cap I_j \neq \emptyset \text{ for strictly more than } 1/2 \text{ the } j \in \{1, 2, ..., m\}\}$$

Then $\mathcal{A}$ is indeed d-close. Fix $\varepsilon > 0$ and suppose $k > (1+\varepsilon) \cdot (1 - \left(\frac{1}{2}\right)^{\frac{1}{d+1}}) \cdot n$, then we can lower bound $|\mathcal{A}|$ as follows. Let A be a uniformly random element of $[n]^{(k)}$, and let X be the number of $j \in \{1, 2, ..., m\}$ such that $A \cap I_j \neq \emptyset$. Then

$$\mathbb{E}[X] = m \cdot \mathbb{P}(A \cap I_1 \neq \emptyset) = m \cdot \left(1 - \frac{\binom{n-d-1}{k}}{\binom{n}{k}}\right).$$

Now

$$\begin{aligned}
\frac{\binom{n-d-1}{k}}{\binom{n}{k}} = \frac{(n-k)(n-k-1)...(n-d-k)}{n(n-1)...(n-d)} &\leq \left(\frac{n-k}{n-d}\right)^{d+1} \\
&< \left(\frac{n-k}{n}\right)^{d+1} \cdot (1 + o_n(1)) \\
&< \left(2^{-\frac{1}{d+1}} - \varepsilon \cdot (1 - 2^{-\frac{1}{d+1}})\right)^{d+1} \cdot (1 + o_n(1)) \\
&< \frac{1}{2},
\end{aligned}$$

for sufficiently large n, depending only on ε. Hence, for sufficiently large n, $\mathbb{E}[X] > \frac{m}{2}$. Now as X is a sum of independent Bernoulli random variables, standard Chernoff bounds

imply X is concentrated about $\mathbb{E}[X]$, so

$$\frac{|\mathcal{A}|}{\binom{n}{k}} = \mathbb{P}(A \in \mathbb{A}) = \mathbb{P}(X > m/2) = 1 - o_n(1),$$

i.e., if $k > (1+\varepsilon) \cdot \left(1 - \left(\frac{1}{2}\right)^{\frac{1}{d+1}}\right) \cdot n$, then $|\mathcal{A}| \geq (1 - o_n(1))\binom{n}{k}$. Since it is clear that $|\mathcal{A}| \leq \binom{n}{k}$, the above shows that this upper bound is essentially tight. We may therefore assume for the rest of the chapter that $k \leq \left(1 - \frac{1}{2}^{\frac{1}{d+1}}\right) \cdot n$.

Since the completion of this thesis the author has been made aware of results for *G-intersecting families*: for a graph G on ground set $[n]$ a set system $\mathcal{A} \subseteq \mathcal{P}([n])$ is said to be G-intersecting if for all $A, B \in \mathcal{A}$ there exist $a \in A$ and $b \in B$ such that $a = b$ or $\{a, b\}$ is an edge of G. The case of d-intersecting families corresponds to taking G to be the dth power of a cycle graph on the ground set. The state of the art restults for G-intersecting families are due to Bohman and Martin [14], where they prove the following result.

Theorem 4.1.1 ([14])**.** *Let G be a graph on $[n]$ with maximum degree Δ and clique number ω. There exists a constant C (depending only on Δ and ω) such that if $\mathcal{H}$ is a G-intersecting k-uniform hypergraph and $k < Cn^{1/2}$ then*

$$|\mathcal{H}| \leq \binom{n}{k} - \binom{n-\omega}{k} + \binom{\omega(\Delta - \omega + 1)}{2}\binom{n-\omega-2}{k-2}.$$

Furthermore, if $\mathcal{H}$ is a G-intersecting family of maximum cardinality then there exists a maximum clique K in G such that $\mathcal{H}$ contains all k-sets that intersect K.

This is a generalisation of our results for $k < Cn^{1/2}$, since we only deal with the case of G being the dth power of a cycle on $[n]$. However we also note that the following conjecture of Bohman, Frieze, Ruszinkó and Thoma is left open in [14]:

Conjecture 4.1.1 (Bohman, Fireze, Ruszinkó and Thoma)**.** *Let*

$$N(C_n, k) = \max\{|\mathcal{H}| : \mathcal{H} \text{ is } C_n\text{-intersecting } k\text{-uniform hypergraph}\}.$$

Then there exists a constant c such that for any fixed $\varepsilon > 0$:

$$k \leq (c-\varepsilon)n \Rightarrow N(C_n,k) = \binom{n}{k} - \binom{n-2}{k} + \binom{n-4}{k-2},$$
$$k \geq (c+\varepsilon)n \Rightarrow N(C_n,k) = (1-o(1))\binom{n}{k}.$$

The main result of this chapter has as an easy corollary the existence of a constant C such that if n is a positive integer and $0 \leq k < \frac{n}{C}$ an integer and $\mathcal{F}_1, \mathcal{F}_2 \subseteq [n]^{(k)}$ are cross 1-close, then

$$\min_{i=1,2} |\mathcal{F}_i| \leq \binom{n}{k} - \binom{n-3}{k} - \binom{n-4}{k-1},$$

with equality if and only if both $\mathcal{F}_1$ and $\mathcal{F}_2$ are identical families isomorphic to

$$\{A \in [n]^{(k)} : \exists a, b \in A \text{ such that } \mathrm{dist}(a,1) \leq 1 \text{ and } \mathrm{dist}(b,2) \leq 1\}.$$

This is a strengthening of Theorem 4.1.1 in the case of $G = C_n$ the cycle of length n (since it improves the constraint $k < O(n^{1/2})$ to $k < O(n)$, and in fact our main result strengthens Theorem 4.1.1 in the same way for G any dth power of a cycle of length n. Furthermore, our main result establishes the existence of a constant c such that if $k \leq cn$ then $N(C_n,k) = \binom{n}{k} - \binom{n-2}{k} + \binom{n-4}{k-2}$, and we have seen that if $k > \left(1 - \sqrt{\frac{1}{2}}\right) \cdot n$ then $N(C_n,k) = (1-o(1))\binom{n}{k}$. Sadly this doesn't establish the whole of Conjecture 4.1.1 since it remains to show that the transition between extremal behaviours of $N(C_n,k)$ as for $k = O(n)$ is sharp around $k = cn$ for some constant c, which remains to be determined.

4.1.1 Extremal families

For a set $A \subseteq [n]$ we define the *d-neighbourhood* of A by

$$D_d(A) := \{b \in [n] \ : \ \exists a \in A \text{ such that } \mathrm{dist}(a,b) \leq d\},$$

and note that a pair of sets $A, B \subseteq [n]$ is d-close if and only if $B \cap D_d(A) \neq \emptyset$. We define for $e \in [n]$ and non-negative integers d, k the following k-uniform family:

$$\mathcal{U}_{e,d,k} = \{A \in [n]^{(k)} \ : \{e, e+1, ..., e+d\} \subseteq D_d(A)\},$$

and we can see that $\mathcal{U}_{e,d,k}$ is d-close. Indeed, suppose $A, B \in \mathcal{U}_{e,d,k}$. Then we have three possible cases as follows:

1. $A \cap \{e, e+1, ..., e+d\}, B \cap \{e, e+1, ..., e+d\} \neq \emptyset$, from which we immediately see the pair $\{A, B\}$ is d-close.

2. $A \cap \{e, e+1, ..., e+d\} = \emptyset$ and $b \in B \cap \{e, e+1, ..., e+d\} \neq \emptyset$. Then since $b \in \{e, e+1, ..., e+d\} \subseteq D_d(A)$ we see that $\{A, B\}$ is d-close. The case with A, B swapped is identical.

3. $A \cap \{e, e+1, ..., e+d\}, B \cap \{e, e+1, ..., e+d\} = \emptyset$. We then take $a \in A$ and $b \in B$ which minimise $\text{dist}(\cdot, e)$ over A and B respectively. Since

$$\{e, e+1, ..., e+d\} \subseteq D_d(A) \cap D_d(B),$$

we must have that $a, b < e \leq a + d, b + d$ (here we have slightly abused notation, and $<$ represents the cyclic order locally around e). Rearranging, this means $e - d \leq a, b < e$, and so $\text{dist}(a, b) \leq d$. Hence the pair A, B is d-close.

In the main theorem we will establish that the families $\mathcal{U}_{e,d,k}$ are extremal uniform d-close families, but first we calculate $|\mathcal{U}_{e,d,k}|$.

Claim 4.1.1. *For positive integer n, for non-negative integers d, k and for $e \in [n]$ we have*

$$|\mathcal{U}_{e,d,k}| = \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1}.$$

Proof. Fix integers n, d, k and $e \in [n]$. Note that

$$\mathcal{U}_{e,d,k} = \{A \in \mathcal{U}_{e,d,k} : A \cap \{e, e+1, ..., e+d\} \neq \emptyset\} \sqcup \{A \in \mathcal{U}_{e,d,k} : A \cap \{e, e+1, ..., e+d\} = \emptyset\},$$

and note that $|\{A \in \mathcal{U}_{e,d,k} : A \cap \{e, e+1, ..., e+d\} \neq \emptyset\}| = \binom{n}{k} - \binom{n-d-1}{k}$.

Now let $S := \{A \in \mathcal{U}_{e,d,k} : A \cap \{e, e+1, ..., e+d\} = \emptyset\}$. Observe that

$$S = \bigsqcup_{i=1}^{d} \bigsqcup_{j=1}^{d+1-i} \{A \in [n]^{(k)} : e-i, e+d+j \in A, A \cap \{e-i+1, ..., e+d+j-1\} = \emptyset\},$$

then since

$$|\{A \in [n]^{(k)} : e-i, e+d+j \in A, A \cap \{e-i+1, ..., e+d+j-1\} = \emptyset\}| = \binom{n-d-j-i-1}{k-2},$$

we have

$$\begin{aligned} |S| &= \sum_{i=1}^{d} \sum_{j=1}^{d+1-i} \binom{n-d-j-i-1}{k-2} \\ &= \sum_{i=1}^{d} \left[\binom{n-d-i-1}{k-1} - \binom{n-2d-2}{k-1} \right] \\ &= \binom{n-d-1}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1}. \end{aligned}$$

So, in total

$$|\mathcal{U}_{e,d,k}| = \binom{n}{k} - \binom{n-d-1}{k} + |S| = \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1},$$

as claimed. □

We also define here, for each $e \in [n]$ and non-negative integer d, the collection

$$\mathcal{P}_{e,d} := \{\{a, b\} \subset ([n] \setminus \{e, e+1, ..., e+d\}) : \{e, e+1, ..., e+d\} \subset D_d(\{a, b\})\},$$

i.e., the pairs of elements disjoint from cyclic interval $\{e, e+1, ..., e+d\}$ whose d-neighbourhood covers the cyclic interval. Note that $|\mathcal{P}_{e,d}| = \frac{d(d+1)}{2}$: indeed for $i = 0, ..., d-1$ we see that in $\mathcal{P}_{e,d}$, the element $e-d+i$ lies in a pair with each of $e+d+1, e+d+2, ..., e+d+1+i$, so there are $\sum_{i=0}^{d-1}(i+1) = \frac{d(d+1)}{2}$ pairs in $\mathcal{P}_{e,d}$ (for n sufficiently large with respect to d).

We are now ready to state the main theorem of this chapter.

Theorem 4.1.2. *Let d be a non-negative integer, then there exists a constant $C = C(d)$, such that the following holds. Let n be a positive integer and let $0 \leq k < n/C$ be an integer, and $\mathcal{F}_1, \mathcal{F}_2 \subseteq [n]^{(k)}$ be cross d-close. Then*

$$\min_{i=1,2} |\mathcal{F}_i| \leq \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1},$$

with equality if and only if there exists an e such that $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{U}_{e,d,k}$.

We get the following immediate corollary.

Corollary 4.1.1. *Let d be a non-negative integer, then there exists a constant $C = C(d)$, such that the following holds. Let n be a positive integer and let $0 \leq k < n/C$ be an integer, and $\mathcal{F} \subseteq [n]^{(k)}$ be d-close. Then*

$$|\mathcal{F}| \leq \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1},$$

with equality if and only if there exists an e such that $\mathcal{F} = \mathcal{U}_{e,d,k}$.

The proof of the main theorem will be given in Section 4.3, but first in Section 4.2 we establish some preliminary theory.

4.2 Preliminaries

4.2.1 Measures and coupling

Our proof will make use of several different *measures*, particularly measures over $\mathcal{P}([n])$ and measures over $S^{(k)}$ (given a set S). In all cases we will define these measures pointwise, i.e., assign measure to each element of $\mathcal{P}([n])$ or $S^{(k)}$ respectively, and then take the measure of a subset to be the sum of the measures of its elements. So for example a measure η on $\mathcal{P}([n])$ will be defined for each singleton $\{A\}$ where $A \subseteq [n]$, and then this definition extended to $\mathcal{A} \subseteq \mathcal{P}([n])$ by $\eta(\mathcal{A}) = \sum_{A\in\mathcal{A}} \eta(\{A\})$.

Recall that a family $\mathcal{F} \subseteq \mathcal{P}([n])$ is called an *upset* if $A \subseteq B \subseteq [n]$ then $A \in \mathcal{F}$ implies $B \in \mathcal{F}$. For a family $\mathcal{F} \subseteq \mathcal{P}([n])$ we will define $\mathcal{F}^{\uparrow}$ to be the smallest upset containing $\mathcal{F}$, i.e., the intersection of all the upsets which contain $\mathcal{F}$. We also recall here the notion of *stochastic domination*: let η_1 and η_2 be measures defined on $\mathcal{P}([n])$, then we say η_1 *stochastically dominates* η_2 if for every upset $\mathcal{F} \subseteq \mathcal{P}([n])$ we have $\eta_1(\mathcal{F}) \geq \eta_2(\mathcal{F})$ (where again $\eta_i(\mathcal{F}) = \sum_{A\in\mathcal{F}} \eta_i(\{A\})$). We write this as $\eta_2 \preceq \eta_1$.

For the first important measure, let $p \in (0,1)$ and then define μ_p to be the *p-biased measure* on $\mathcal{P}([n])$, i.e., for $A \subseteq [n]$, we set $\mu_p(\{A\}) = p^{|A|}(1-p)^{n-|A|}$, and for $\mathcal{A} \subseteq \mathcal{P}([n])$, we let $\mu_p(\mathcal{A}) = \sum_{A\in\mathcal{A}} \mu_p(\{A\})$.

For the second important measure, let S be a set and $\mathcal{A} \subseteq S^{(k)}$. We write $\mu(\mathcal{A})$ for the *uniform measure* $|\mathcal{A}|/\binom{|S|}{k}$.

We also define the following less standard measure and coupled random variables that are critical to our proof. Fix positive integer n and non-negative integer d, then let

$$f_d : \{0,1,2\}^n \to \{0,1\}^n; \quad f_d(X)_i = \begin{cases} 1 & \text{if } X_i = 1 \text{ and } X_j \neq 2 \text{ for } \mathrm{dist}(i,j) \leq d, \\ 0 & \text{otherwise,} \end{cases}$$

and similarly

$$g_d : \{0,1,2\}^n \to \{0,1\}^n; \quad g_d(X)_i = \begin{cases} 1 & \text{if } X_i = 2 \text{ and } X_j \neq 1 \text{ for } \operatorname{dist}(i,j) \leq d, \\ 0 & \text{otherwise,} \end{cases}$$

and note that for all $X \in \{0,1,2\}^n$, the coupled random variables $f_d(X)$ and $g_d(X)$ are *not* d-close. Letting X be uniformly distributed in $\{0,1,2\}^n$ we define the measure

$$\nu_d : \mathcal{P}(\mathcal{P}([n])) \to \mathbb{R}_{\geq 0}; \quad \nu_d(\mathcal{A}) = \mathbb{P}(f_d(X) \in \mathcal{A}).$$

Key Observation: Firstly, for $\mathcal{A} \subseteq \mathcal{P}([n])$ it is clear that $\nu_d(\mathcal{A}) = \mathbb{P}(g_d(X) \in \mathcal{A})$. Hence, if $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}([n])$ are cross d-close, then

$$\begin{aligned} \nu_d(\mathcal{A}) + \nu_d(\mathcal{B}) &= \mathbb{P}(f_d(X) \in \mathcal{A}) + \mathbb{P}(g_d(X) \in \mathcal{B}) \\ &= \mathbb{P}(\{f_d(X) \in \mathcal{A}\} \cup \{g_d(X) \in \mathcal{B}\}) \leq 1, \end{aligned}$$

where the second equality follows since $\mathcal{A}, \mathcal{B}$ being cross d-close implies the events $\{f_d(X) \in \mathcal{A}\}$ and $\{g_d(X) \in \mathcal{B}\}$ are disjoint.

4.2.2 Relations between measures

We will require the following proposition later.

Proposition 4.2.1. *Let d be a non-negative integer, and let $0 < p \leq \frac{1}{3^{2d+3}+1}$. Then ν_d stochastically dominates μ_p.*

The proof of this result requires the following theorem due to Holley [45] (which is itself a generalisation of the Fortuin-Kasteleyn-Ginibre inequality [30]).

Theorem 4.2.1 (Holley's Theorem)**.** *Let $\eta_1, \eta_2 : \mathcal{P}(\mathcal{P}([n])) \to \mathbb{R}$ be measures satisfying*

$$\eta_1(x \cup y)\eta_2(x \cap y) \geq \eta_1(x)\eta_2(y),$$

for all $x, y \in \mathcal{P}([n])$. Then η_1 stochastically dominates η_2.

Proof of Proposition 4.2.1. Here we identify $\mathcal{P}([n])$ with $\{0,1\}^n$ in the natural way ($A \subseteq [n] \leftrightarrow (x_i)_{i=1}^n : x_i = 1 \iff i \in A$). We will show that for $0 < p \leq \frac{1}{3^{2d+3}+1}$ we have

$$\nu_d(x \cup y)\mu_p(x \cap y) \geq \nu_d(x)\mu_p(y)$$

for all $x, y \in \mathcal{P}([n])$, from which Holley's Theorem implies the result. Rearranging the above equation, it is equivalent to

$$\frac{\nu_d(x \cup y)}{\nu_d(x)} \geq \frac{\mu_p(y)}{\mu_p(x \cap y)} = \left(\frac{p}{1-p}\right)^{|y \setminus x|}.$$

We claim that $\frac{\nu_d(x \cup \{z\})}{\nu_d(x)} \geq \frac{p}{1-p}$ for all $z \notin x$. The above inequality then follows as

$$\frac{\nu_d(x \cup y)}{\nu_d(x)} = \frac{\nu_d(x \cup \{y_1, ..., y_k\})}{\nu_d(x)} = \prod_{i=1}^{k} \frac{\nu_d(x \cup \{y_1, ..., y_i\})}{\nu_d(x \cup \{y_1, ..., y_{i-1}\})} \geq \left(\frac{p}{1-p}\right)^{|y \setminus x|},$$

where $\{y_1, ..., y_k\} = y \setminus x$. It remains to prove the claim.

Recalling the identification between $\mathcal{P}([n])$ and $\{0,1\}^n$, the claim is equivalent to showing

$$\frac{\nu_d(X_1X_2...X_{i-1}1X_{i+1}...X_n)}{\nu_d(X_1X_2...X_{i-1}0X_{i+1}...X_n)} \geq \frac{p}{1-p}$$

for each $i \in [n]$ and sequence of bits $X_1, X_2, ..., X_{i-1}, X_{i+1}, ..., X_n$. We write $\mathbf{X}_+ = X_1X_2...X_{i-1}1X_{i+1}...X_n$ and $\mathbf{X}_- = X_1X_2...X_{i-1}0X_{i+1}...X_n$.

We now fix $i \in [n]$ and the bits $X_1, X_2, ..., X_{i-1}, X_{i+1}, ..., X_n$. Let L be the number of consecutive 0's to the left of X_i and R the number of consecutive 0's to the right of X_i, i.e.,

$$X_1X_2...X_{i-1}X_iX_{i+1}...X_n = X_1...X_{i-L-1}\underbrace{0...0}_{L}X_i\underbrace{0...0}_{R}X_{i+R+1}...X_n,$$

X_i is 1 or 0 respectively, either $X_{i-L-1} = 1$ or $i = L + 1$ and either $X_{i+R+1} = 1$ or

$i + R = n$. We prove the claim for each possible combination of values L, R, and split into the following cases:

- **Case 1:** $L \leq i-2, R \leq n-i-1$,
- **Case 2:** $L = i-1, R \leq n-i-1$ (the case $L \leq i-2, R = n-i$ follows symmetrically),
- **Case 3:** $L = i-1, R = n-i$.

We only prove Case 1 in detail, the remaining cases follow very similarly.

Proof of Case 1. In the first case, we define

$$U = \left\{ 1\mathbf{Z_l}1\mathbf{Z_r}1 \in \{0,1,2\}^{L+R+3} \ : \ \begin{array}{l} \mathbf{Z_l} \in \{0,1,2\}^L, \mathbf{Z_r} \in \{0,1,2\}^R, \\ f_d(1\mathbf{Z_l}1\mathbf{Z_r}1) = 10...010...01 \end{array} \right\}$$

and

$$V = \left\{ 1\mathbf{Z_l}Z_{L+1}\mathbf{Z_r}1 \in \{0,1,2\}^{L+R+3} \ : \ \begin{array}{l} \mathbf{Z_l} \in \{0,1,2\}^L, \mathbf{Z_r} \in \{0,1,2\}^R, \\ f_d(1\mathbf{Z_l}Z_{L+1}\mathbf{Z_r}1)10...000...01 \end{array} \right\}.$$

We note that

$$\frac{\nu_d(\mathbf{X}_+)}{\nu_d(\mathbf{X}_-)} = \frac{|U|}{|V|}.$$

Indeed, this follows by considering a bipartite graph with vertex sets

$$A = \{\mathbf{W} \in \{0,1,2\}^n : f_d(\mathbf{W}) = \mathbf{X}_+\} \text{ and } B = \{\mathbf{W}' \in \{0,1,2\}^n : f_d(\mathbf{W}') = \mathbf{X}_-\},$$

and an edge from $\mathbf{W} \in A$ to $\mathbf{W}' \in B$ when $\mathbf{W} = W_1W_2...W_n, \mathbf{W}' = W'_1W'_2...W'_n$ satisfy both

$$W_1W_2...W_{i-L-2} = W'_1W'_2...W'_{i-L-2}$$

and

$$W_{i+R+2}W_{i+R+3}...W_n = W'_{i+R+2}W'_{i+R+3}...W'_n.$$

Then $\mathbf{W} \in A$ is joined precisely to the collection

$$\{W_1...W_{i-L-2}\mathbf{Z}W_{i+R+2}...W_n : \mathbf{Z} \in V\} \subset B,$$

and $\mathbf{W}' \in B$ is joined precisely to the collection

$$\{W'_1...W'_{i-L-2}\mathbf{Z}W'_{i+R+2}...W'_n : \mathbf{Z} \in U\} \subset A.$$

(Note that $\nu_d(\mathbf{X}_+) = \frac{|A|}{3^n}$ and $\nu_d(\mathbf{X}_-) = \frac{|B|}{3^n}$.) Hence the degree of every $\mathbf{W} \in A$ is $|V|$, and the degree of every $\mathbf{W}' \in B$ is $|U|$ and by double counting edges of the bipartite graph we have $|A||V| = |B||U|$. It follows that

$$\frac{\nu_d(\mathbf{X}_+)}{\nu_d(\mathbf{X}_-)} = \frac{|A|}{|B|} = \frac{|U|}{|V|}.$$

Thus the claim is equivalent to

$$\frac{|U|}{|V|} \geq \frac{p}{1-p}.$$

Suppose first that $L, R \geq d+1$. We define a second bipartite graph with vertex classes U, V and join $1Z_1...Z_LZ_{L+1}Z_{L+2}...Z_{L+R+1}1 \in V$ to each of

$$1Z_1...Z_{L-d-1}2Y_1...Y_d1Y_{d+1}...Y_{2d}2Z_{L+d+3}...Z_{L+R+1}1 \in U$$

where each $Y_j \in \{0,1\}$. Each element of V has degree 2^{2d} and each element of U has degree at most 3^{2d+3}, and by double counting edges we can see that

$$\frac{|U|}{|V|} \geq \left(\frac{2}{3}\right)^{2d}\frac{1}{27} \geq \frac{1}{3^{2d+3}}.$$

Suppose now that $L, R \leq d$. Then, with U and V as before, we note that $U =$

$\{1\underbrace{0...0}_{L}1\underbrace{0...0}_{R}1\}$, so has size 1, while V has size at most $3^{L+R+1} \leq 3^{2d+3}$. Hence $\frac{|U|}{|V|} \geq \frac{1}{3^{2d+3}}$.

Finally if $L \leq d, R \geq d+1$, similar reasoning will give $\frac{|U|}{|V|} \geq \frac{2^d}{3^{L+d+2}} \geq \frac{1}{3^{2d+3}}$.

Since $p \leq \frac{1}{3^{2d+3}+1}$ we have $\frac{p}{1-p} \leq \frac{1}{3^{2d+3}}$, and see that for all values of L and R satisfying the conditions of the first case, we have $\frac{|U|}{|V|} \geq \frac{p}{1-p}$. Hence ν_d stochastically dominates μ_p, as required, and the result follows.

□

Proof of Case 2. In the second case, we define

$$U = \left\{ \mathbf{Z_l}1\mathbf{Z_r}1 \in \{0,1,2\}^{L+R+2} \ : \begin{array}{l} \mathbf{Z_l} \in \{0,1,2\}^L, \mathbf{Z_r} \in \{0,1,2\}^R, \\ f_d(\mathbf{Z_l}1\mathbf{Z_r}1) = 0...010...01 \end{array} \right\}$$

and

$$V = \left\{ \mathbf{Z_l}Z_{L+1}\mathbf{Z_r}1 \in \{0,1,2\}^{L+R+2} \ : \begin{array}{l} \mathbf{Z_l} \in \{0,1,2\}^L, \mathbf{Z_r} \in \{0,1,2\}^R, \\ f_d(\mathbf{Z_l}Z_{L+1}\mathbf{Z_r}1) = 0...000...01 \end{array} \right\},$$

and similar reasoning to the first case proves that

$$\frac{|U|}{|V|} \geq \frac{1}{3^{2d+3}}.$$

Since $p \leq \frac{1}{3^{2d+3}+1}$, we see that $\frac{|U|}{|V|} \geq \frac{p}{1-p}$, and the result follows. □

Proof of Case 3. In the final case, we define

$$U = \left\{ \mathbf{Z_l}1\mathbf{Z_r} \in \{0,1,2\}^{L+R+1} \ : \begin{array}{l} \mathbf{Z_l} \in \{0,1,2\}^L, \mathbf{Z_r} \in \{0,1,2\}^R, \\ f_d(Z_1...Z_L1Z_{L+2}...Z_{L+R+1}) = 0...010...0 \end{array} \right\}$$

and

$$V = \left\{ \mathbf{Z_l} Z_{L+1} \mathbf{Z_r} \in \{0,1,2\}^{L+R+1} \ : \ \begin{array}{l} \mathbf{Z_l} \in \{0,1,2\}^L, \mathbf{Z_r} \in \{0,1,2\}^R, \\ f_d(Z_1 ... Z_L Z_{L+1} Z_{L+2} ... Z_{L+R+1}) = 0...000...0 \end{array} \right\},$$

and similar reasoning to the first case proves that

$$\frac{|U|}{|V|} \geq \frac{1}{3^{2d+3}}.$$

Since $p \leq \frac{1}{3^{2d+3}+1}$, we see that $\frac{|U|}{|V|} \geq \frac{p}{1-p}$, and the result follows. □

□

4.2.3 Juntas

In our argument it is important to be able to define subfamilies using intersection conditions. For this we use the notion of *slices*: let $\mathcal{F} \subseteq \mathcal{P}([n])$ be a family, let $S \subseteq [n]$ be a set, then S gives rise to a partition of $\mathcal{F}$ into $2^{|S|}$ *slices*, $\{\mathcal{F}_S^B\}_{B \subseteq S}$, where parts are determined by intersection with S:

$$\mathcal{F}_S^B = \{A \setminus B : (A \in \mathcal{F}) \wedge (A \cap S = B)\}.$$

We will consider families $\mathcal{F}_S^B$ as lying in $([n] \setminus S)^{(k-|B|)}$.

The proof of Theorem 4.1.2 is an adaptation of the 'Junta Method' which was introduced to Extremal Combinatorics in through the work of Dinur and Friedgut [23] who were motivated by results in the analysis of Boolean functions, and was recently applied by Keller and Lifshitz [54] to the Erdős-Chvátal conjecture. Here we give an introduction to the notion of a *junta*.

Definition 4.2.1. *Let $n, k \in \mathbb{N}$, and let $J \subseteq [n]$ satisfy $|J| < k$. A set system $\mathcal{J} \subseteq [n]^{(k)}$ is called a J-junta if $A, B \in [n]^{(k)}$ are such that $A \cap J = B \cap J$ then $A \in \mathcal{J}$ if and only if $B \in \mathcal{J}$ (i.e., inclusion in $\mathcal{J}$ is completely determined by intersection with J). A j-junta*

is a J-junta for some set J of size j.

The notion of junta has its origins in the context of Boolean functions. A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is called a junta if it is only dependent on a small number of the coordinates. Juntas are therefore a natural notion for 'simple' Boolean functions, and there has been extensive research into approximating Boolean functions by juntas, the earliest of which is the Junta theorem of Friedgut [31] which shows that a Boolean function with small 'total influence' (which is precisely the edge boundary of the the vertex set $f^{-1}(1) \subseteq \{0,1\}^n = Q_n$) is approximated by a junta depending on a constant number of coordinates. Ideas from the analysis of Boolean functions have seen applications in extremal combinatorics, particularly for example in Erdős-Ko-Rado type theorems [23, 25, 26, 33], but also in isoperimetry [51].

Approximation by junta approaches have been appplied in the proofs of important results in the analysis of Boolean functions, such as [32, 43]. More generally, the analysis of Boolean functions has very wide ranging applications beyond combinatorics, especially to theoretical computer science in such areas as algorithms research (e.g. computational learning theory [60, 72], property testing [13, 35, 67]) and machine learning [64]) but also to areas such as the theory of social choice [9, 52], and cryptography [36]. A survey of the methods and applications of the analysis of Boolean functions is in the book by O'Donnell [63].

For a set $J \subseteq [n]$ (typically J will be of 'constant' size), a family $\mathcal{J} \subseteq \mathcal{P}(J)$ and $k \geq |J|$, the *k-uniform junta* generated by $\mathcal{J}$ is

$$\langle \mathcal{J} \rangle = \left\{ A \in [n]^{(k)} : A \cap J \in \mathcal{J} \right\}.$$

Another natural k-uniform junta that arises from $\mathcal{J}$ is $\left(\mathcal{J}^{\uparrow}\right)^{(k)}$, i.e., the intersection of $\mathcal{J}^{\uparrow}$, the smallest upset contained in $\mathcal{P}([n])$ containing $\mathcal{J}$, and $[n]^{(k)}$. We have $\langle \mathcal{J} \rangle \subseteq \left(\mathcal{J}^{\uparrow}\right)^{(k)}$, but the following lemma by Keller and Lifshitz [54] shows that these juntas are essentially

the same in the sense that

$$\mu\left(\left(\mathcal{J}^{\uparrow}\right)^{(k)} \setminus \langle\mathcal{J}\rangle\right) = o(\mu\left(\langle\mathcal{J}\rangle\right)).$$

We provide their proof here for completeness.

Lemma 4.2.1. *Let $J \subseteq [n]$, where $j = |J|$ is a constant. Let $\mathcal{J} \subseteq \mathcal{P}(J)$ be a family, and let $l < k$ be the minimal size of an element of $\mathcal{J}$. Then*

$$\left|\left(\mathcal{J}^{\uparrow}\right)^{(k)}\right| = \left|\mathcal{J}^{(l)}\right| \binom{n}{k-l} + O_j\left(\left(\frac{k}{n}\right)^{l+1}\binom{n}{k}\right), \tag{4.1}$$

and

$$\mu\left(\left(\mathcal{J}^{\uparrow}\right)^{(k)} \setminus \langle\mathcal{J}\rangle\right) = O_j\left(\left(\frac{k}{n}\right)^{l+1}\right). \tag{4.2}$$

Proof. For (4.1), we note that

$$\left|\left(\mathcal{J}^{\uparrow}\right)^{(k)}\right| = \sum_{\{A\subseteq J:|A|\geq l+1\}}\left|\left(\left(\mathcal{J}^{\uparrow}\right)^{(k)}\right)_J^A\right| + \sum_{\{A\subseteq J:|A|=l\}}\left|\left(\left(\mathcal{J}^{\uparrow}\right)^{(k)}\right)_J^A\right|.$$

The lemma then follows from the following inequalities:

$$\sum_{\{A\subseteq J:|A|\geq l+1\}}\left|\left(\left(\mathcal{J}^{\uparrow}\right)^{(k)}\right)_J^A\right| \leq \sum_{\{A\subseteq J:|A|\geq l+1\}}\binom{n-j}{k-|A|} = O_j\left(\left(\frac{k}{n}\right)^{l+1}\binom{n}{k}\right),$$

and

$$\sum_{\{A\subseteq J:|A|=l\}}\left|\left(\left(\mathcal{J}^{\uparrow}\right)^{(k)}\right)_J^A\right| = \left|\mathcal{J}^{(l)}\right|\binom{n-j}{k-l} = \left|\mathcal{J}^{(l)}\right|\binom{n}{k-l} + O_j\left(\left(\frac{k}{n}\right)^{l+1}\binom{n}{k}\right).$$

We see that (4.2) follows from noting that

$$|\langle\mathcal{J}\rangle| = \sum_{A\in\mathcal{J}} \binom{n-j}{k-|A|} = |\mathcal{J}^{(l)}|\binom{n-j}{k-l} + O_j\left(\left(\frac{k}{n}\right)^{l+1}\binom{n}{k}\right)$$
$$= \left|\mathcal{J}^{(l)}\right|\binom{n}{k-l} + O_j\left(\left(\frac{k}{n}\right)^{l+1}\binom{n}{k}\right),$$

and so, combining this with (4.1) we have

$$|\left(\mathcal{J}^{\uparrow}\right)^{(k)} \setminus \langle\mathcal{J}\rangle| = O_j\left(\left(\frac{k}{n}\right)^{l+1}\binom{n}{k}\right),$$

from which (4.2) immediately follows. □

Roughly speaking, the approach to proving Theorem 4.1.2 is as follows: we first demonstrate that if $\mathcal{F}_1, \mathcal{F}_2 \subseteq [n]^{(k)}$ are a cross d-close pair of families then there is a pair of k-uniform juntas $\mathcal{G}_1, \mathcal{G}_2$ that is also cross d-close, and such that each family $\mathcal{F}_i$ is essentially contained in the junta $\mathcal{G}_i$, in an appropriately defined sense. We next show that if $\mathcal{G}_1, \mathcal{G}_2$ is a cross d-close pair of k-uniform juntas, then $\min_{i\in\{1,2\}} |\mathcal{G}_i| \leq |\mathcal{U}_{e,d,k}|$. Furthermore, we show that if the juntas have nearly maximum size, in an appropriate sense, then the juntas $\mathcal{G}_1, \mathcal{G}_2$ are contained in the same $\mathcal{U}_{e,d,k}$. Having completed these two steps, we can see that if families $\mathcal{F}_1, \mathcal{F}_2$ are k-uniform, cross d-close and $\min_{i\in\{1,2\}} |\mathcal{F}_i| \geq |\mathcal{U}_{e,d,k}|$, then each of the families $\mathcal{F}_i$ is a small alteration of the same $\mathcal{U}_{e,d,k}$. A final step strengthens this stability result to an exact result using a bootstrapping lemma.

In order to identify the 'junta part' of families, we will require the following definition and lemma introduced by Dinur and Friedgut [23], and a proposition by Keller and Lifshitz [54].

Definition 4.2.2. *Let s be a non-negative integer, and let $\varepsilon \in (0,1)$. A family $\mathcal{F} \subseteq [n]^{(k)}$ is called* (s,ε)-capturable *if there exists a set $S \subseteq [n]$ of size at most s, such that $\mu(\mathcal{F}_S^{\emptyset}) \leq \varepsilon$. Otherwise we say the family is* (s,ε)-uncapturable. *For intuition, this definition says that a k-uniform family $\mathcal{F}$ is (s,ε)-capturable if there exists a set $S \subseteq [n]$*

of size at most s such that the number of elements of $\mathcal{F}$ that are disjoint from S is at most an ε-proportion of $([n] \setminus S)^{(k)}$ (i.e., it is small).

Lemma 4.2.2. *For any constants $\zeta \in (0, \frac{1}{2})$ and $r \in \mathbb{N}$, there exists a constant $s = s(\zeta, r)$ such that the following holds. Let $n, k \in \mathbb{N}$ and $p \in (\zeta, 1)$ be numbers such that $\frac{k}{n} \leq \frac{p}{2}$, and let $\mathcal{A} \subseteq [n]^{(k)}$ be a family that satisfies $\mu_p(\mathcal{A}^{\uparrow}) \leq 1 - \zeta$. Then $\mathcal{A}$ is $(s, \left(\frac{k}{n}\right)^r)$-capturable.*

Proposition 4.2.2. *Let $r, s \in \mathbb{N}$ be constants, and let $C = (s+1)^r$. For any $k < n$, for any $\varepsilon \geq \left(\frac{k}{n}\right)^r$, and for any family $\mathcal{F} \subseteq [n]^{(k)}$, there exists a set $J \subseteq [n]$ of size C and a family $\mathcal{J} \subset \mathcal{P}(J)$, such that:*

1. *For each $B \in \mathcal{J}$, the family $\mathcal{F}_B^B$ is $\left(s, \varepsilon \left(\frac{n}{k}\right)^{|B|}\right)$-uncapturable.*
2. *We have*
$$\mu\left(\mathcal{F} \setminus \mathcal{J}^{\uparrow}\right) \leq C\varepsilon.$$

4.3 Proof of the main theorem

4.3.1 Approximating cross d-close pairs of families with juntas

Our aim in this subsection is to find, for a pair of k-uniform, cross d-close families $\mathcal{F}_1, \mathcal{F}_2$, a pair of k-uniform, cross d-close juntas $\mathcal{G}_1, \mathcal{G}_2$ which approximate $\mathcal{F}_1, \mathcal{F}_2$ respectively in the sense that only a very small measure of $\mathcal{F}_i$ lies outside of $\mathcal{G}_i$. We will prove the following theorem.

Theorem 4.3.1. *Let r be constant, let $k < \frac{n}{2} \cdot \frac{1}{3^{2d+3}+1}$, and let $\mathcal{F}_1, \mathcal{F}_2 \subseteq [n]^{(k)}$ be families that are cross d-close. Then there exist $O_r(1)$-juntas $\mathcal{G}_1, \mathcal{G}_2 \subseteq [n]^{(k)}$ that are cross d-close, such that $\mu(\mathcal{F}_i \setminus \mathcal{G}_i) = O_r\left(\left(\frac{k}{n}\right)^r\right)$ for $i = 1, 2$.*

In order to prove Theorem 4.3.1 we require the following proposition, which shows that, provided k_1, k_2 are integers not too large with respect to n, if each $\mathcal{F}_i$ is a k_i-uniform family, and $\mathcal{F}_1, \mathcal{F}_2$ is a cross d-close pair then at least one of the $\mathcal{F}_i$ is capturable using a set of only constant size.

Proposition 4.3.1. *For constant $r \in \mathbb{N}$, there exists $s = s(r)$ such that the following holds. Let $k_1, k_2 < \frac{n}{2} \cdot \frac{1}{3^{2d+3}+1}$, and let $\mathcal{F}_1 \subseteq [n]^{(k_1)}, \mathcal{F}_2 \subseteq [n]^{(k_2)}$ be families that are cross d-close. Then there exists $i \in \{1,2\}$ such that $\mathcal{F}_i$ is $\left(s, \left(\frac{k_i}{n}\right)^r\right)$-capturable.*

Proof. Let $k_1, k_2 < \frac{n}{2} \cdot \frac{1}{3^{2d+3}+1}$, and let $\mathcal{F}_1 \subseteq [n]^{(k_1)}, \mathcal{F}_2 \subseteq [n]^{(k_2)}$ be families that are cross d-close. Then the upsets $\mathcal{F}_1^{\uparrow}, \mathcal{F}_2^{\uparrow}$ are also cross d-close. Let $p = \frac{1}{3^{2d+3}+1}$ By Proposition 4.2.1 we have

$$\mu_p(\mathcal{F}_1^{\uparrow}) + \mu_p(\mathcal{F}_2^{\uparrow}) \leq \nu_d(\mathcal{F}_1^{\uparrow}) + \nu_d(\mathcal{F}_2^{\uparrow}) \leq 1.$$

So there exists $i \in \{1,2\}$ such that $\mu_p(\mathcal{F}_i^{\uparrow}) \leq 1/2$. Applying Lemma 4.2.2 with $\zeta = \frac{p}{2}$ (the assumption $\frac{k}{n} \leq \frac{p}{2}$ holds as $k_i < \frac{n}{2} \cdot \frac{1}{3^{2d+3}+1}$), we find a constant $s = s(r)$ such that the family $\mathcal{F}_i$ is $\left(s, \left(\frac{k_i}{n}\right)^r\right)$-capturable. □

We remark here that in the case k_1, k_2 are small, say constant size, then it is not surprising that some $\mathcal{F}_i$ is capturable with a constant sized set: suppose $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$, then let $S = D_d(A \cup B)$. The set S will have constant size, at most $(2d+1)(k_1+k_2)$, and noting $(\mathcal{F}_1)_S^{\emptyset} = \{X \in \mathcal{F}_1 \ : \ X \cap S = \emptyset\} = \emptyset$ (indeed, if $X \in \mathcal{F}_1$ such that $X \cap S = \emptyset$, then $X \cap D_d(B) = \emptyset$ contradicting the fact that $\mathcal{F}_1$ and $\mathcal{F}_2$ are d-close) we see that $\mathcal{F}_1$ is in fact $(|S|, 0)$-capturable.

We are now ready to prove Theorem 4.3.1

Proof of Theorem 4.3.1. Let $s = s(r)$ be a constant to be defined below. By Proposition 4.2.2 applied with $\varepsilon = (k/n)^r$, there exist sets J_1, J_2 of size $O_r(1)$ each, and families $\mathcal{J}_1 \subseteq \mathcal{P}(J_1), \mathcal{J}_2 \subseteq \mathcal{P}(J_2)$, such that for $i = 1, 2$ we have:

1. For each $B \in \mathcal{J}_i$, the family $(\mathcal{F}_i)_B^B$ is $\left(s, \left(\frac{k}{n}\right)^{r-|B|}\right)$-uncapturable.
2. We have $\mu(\mathcal{F}_i \setminus \mathcal{J}_i^{\uparrow}) = O_r\left(\left(\frac{k}{n}\right)^r\right)$.

By Lemma 4.2.1 we may remove from each $\mathcal{J}_i$ all sets of size at least r, so without loss of generality we may assume that $|B| < r$ for every $B \in \mathcal{J}_i$.

We claim that the families $\mathcal{J}_1^{\uparrow}$ and $\mathcal{J}_2^{\uparrow}$ are cross d-close. Suppose to the contrary there exists $A_1 \in \mathcal{J}_1^{\uparrow}, A_2 \in \mathcal{J}_2^{\uparrow}$ such that A_1 and A_2 are not d-close. Then there exists $B_1 \in \mathcal{J}_1$ such that $B_1 \subseteq A_1$, and $B_2 \in \mathcal{J}_2$ such that $B_2 \subseteq A_2$, and so B_1 and B_2 are not d-close. Let $E = D_d(B_1 \cup B_2)$, and note that

$$|E| \leq (2d+1)(|B_1| + |B_2|) \leq 2(2d+1)(r-1).$$

To find a contradiction we show the following contradicting claims hold.

Claim 4.3.1. *The families* $(\mathcal{F}_1)_E^{B_1}, (\mathcal{F}_2)_E^{B_2}$ *are* $\left(s - 2(2d+1)(r-1), \left(\frac{k}{n}\right)^r\right)$*-uncapturable.*

Proof. Since for $i = 1, 2$ we have $(\mathcal{F}_i)_{B_i}^{B_i}$ is $\left(s, \left(\frac{k}{n}\right)^r\right)$-uncapturable, then $(\mathcal{F}_i)_E^{B_i}$ is $\left(s - |E \setminus B_i|, \left(\frac{k}{n}\right)^r\right)$-uncapturable. Indeed, suppose for a contradiction that $(\mathcal{F}_i)_E^{B_i}$ is $\left(s - |E \setminus B_i|, \left(\frac{k}{n}\right)^r\right)$-capturable, so there exists a set $S \subseteq [n] \setminus E$ of size at most $s - |E \setminus B_i|$ such that $\mu\left(\left((\mathcal{F}_i)_E^{B_i}\right)_S^{\emptyset}\right) \leq (\frac{k}{n})^r$. Letting $S_* = S \cup (E \setminus B_i)$, we have $|S_*| \leq s$ and

$$\left((\mathcal{F}_i)_{B_i}^{B_i}\right)_{S_*}^{\emptyset} = \left((\mathcal{F}_i)_E^{B_i}\right)_S^{\emptyset},$$

and so $(\mathcal{F}_i)_{B_i}^{B_i}$ is $\left(s, \left(\frac{k}{n}\right)^r\right)$-capturable, a contradiction.

The claim follows by noting that

$$|E \setminus B_i| \leq 2(2d+1)(r-1).$$

□

Claim 4.3.2. *The families* $(\mathcal{F}_1)_E^{B_1}, (\mathcal{F}_2)_E^{B_2}$ *are cross* d*-close.*

Proof. Suppose for a contradiction that $C_1 \in (\mathcal{F}_1)_E^{B_1}$ and $C_2 \in (\mathcal{F}_2)_E^{B_2}$ are not d-close. Then $C_1 \cup B_1 \in \mathcal{F}_1$ and $C_2 \cup B_2 \in \mathcal{F}_2$ are not d-close. Indeed, suppose $a \in C_1 \cup B_1$ and $b \in C_2 \cup B_2$ satisfy $\text{dist}(a, b) \leq d$. As C_1, C_2 are not d-close, we can't have $a \in C_1$ and

$b \in C_2$, and as B_1, B_2 are not d-close we can't have $a \in B_1, b \in B_2$.

Hence either $a \in C_1, b \in B_2$ or $a \in B_1, b \in C_2$. By symmetry we may assume the former, in which case since $b \in B_2$ and $\text{dist}(a,b) \leq d$ we have $a \in E$. But by assumption $C_1 \cap E = \emptyset$, a contradiction.

It follows that $C_1 \cup B_1 \in \mathcal{F}_1$ and $C_2 \cup B_2 \in \mathcal{F}_1$ are not d-close, contradicting our assumption about the d-closeness of $\mathcal{F}_1, \mathcal{F}_2$. Hence $(\mathcal{F}_1)_E^{B_1}, (\mathcal{F}_2)_E^{B_2}$ are cross d-close. □

These claims are contradictory. Indeed, by applying Proposition 4.3.1 to the pair of families $(\mathcal{F}_1)_E^{B_1}, (\mathcal{F}_2)_E^{B_2}$, which are d-close by Claim 4.3.2, there exists an $i \in \{1,2\}$ such that $(\mathcal{F}_i)_E^{B_i}$ is $\left(s', \left(\frac{k}{n}\right)^r\right)$-capturable for some constant s' depending only on r. Taking $s = s' + 2(2d+1)(r-1)$, which is a constant depending only on r for fixed d, we contradict Claim 4.3.1.

Hence the families $\mathcal{J}_1^{\uparrow}, \mathcal{J}_2^{\uparrow}$ are cross d-close and taking $\mathcal{G}_i = \langle \mathcal{J}_i \rangle$ for $i = 1, 2$ proves the theorem. □

4.3.2 Determining the structure of pairs of 'large' juntas that are cross d-close

We now show that if $\mathcal{G}_1, \mathcal{G}_2$ is a pair of k-uniform, cross d-close juntas, then $\min_i |\mathcal{G}_i| \leq |\mathcal{U}_{e,d,k}|$, and furthermore, this bound is nearly tight if and only if there is some e such that $\mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{U}_{e,d,k}$.

First we observe that the cross d-close property for a pair of juntas, say $\langle \mathcal{J}_1 \rangle, \langle \mathcal{J}_2 \rangle$, is inherited by the pair of families of intersections which define those juntas, i.e., $\mathcal{J}_1, \mathcal{J}_2$.

Claim 4.3.3. *For any constant j there exists a constant $C(j)$ such that the following holds. Let $J \in [n]^{(j)}$, let $k \leq n/C$, and let $\langle \mathcal{J}_1 \rangle, \langle \mathcal{J}_2 \rangle \subseteq [n]^{(k)}$ be J-juntas that are cross d-close. Then the families $\mathcal{J}_1, \mathcal{J}_2$ are cross d-close as well.*

Proof. Suppose for a contradiction that $A_1 \in \mathcal{J}_1, A_2 \in \mathcal{J}_2$ are not d-close. Noting that

$|D_d(J)| \leq (2d+1)j$ is also constant size, so long as C is sufficiently large then there exist $B_1 \in ([n] \setminus D_d(J))^{(k-|A_1|)}, B_2 \in ([n] \setminus D_d(J))^{(k-|A_2|)}$ that are not d-close over $[n]$.

Then $A_1 \cup B_1 \in \langle \mathcal{J}_1 \rangle, A_2 \cup B_2 \in \langle \mathcal{J}_2 \rangle$ are also not d-close: indeed suppose there exist $a \in A_1 \cup B_1$ and $b \in A_2 \cup B_2$ such that $\mathrm{dist}(a,b) \leq d$. Then as A_1, A_2 are not d-close we can't have $a \in A_1, b \in A_2$, and as B_1, B_2 are not d-close we can't have $a \in B_1, b \in B_2$, so without loss of generality we may assume that $a \in A_1, b \in B_2$. As $a \in A_1 \subseteq J$ and $\mathrm{dist}(a,b) \leq d$, we have that $b \in D_d(J)$. This contradicts the fact that $B_1 \cap D_d(J) = \emptyset$.

It follows that $A_1 \cup B_1 \in \langle \mathcal{J}_1 \rangle, A_2 \cup B_2 \in \langle \mathcal{J}_2 \rangle$ are not d-close, contradicting the assumption that $\langle \mathcal{J}_1 \rangle, \langle \mathcal{J}_2 \rangle$ are d-close. Thus $\mathcal{J}_1, \mathcal{J}_2$ are d-close as claimed. □

Proposition 4.3.2. *For any constant j, there exists a constant $C(j)$, such that the following holds. Let $J \in [n]^{(j)}$, and let $k \leq n/C$. Suppose that $\mathcal{G}_1, \mathcal{G}_2 \subseteq [n]^{(k)}$ are J-juntas that are cross d-close. Then*

$$\min_{i=1,2} |\mathcal{G}_i| \leq \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1}.$$

Moreover, if

$$\min_{i=1,2} |\mathcal{G}_i| \geq \binom{n}{k} - \binom{n-d-1}{k} + \left(\frac{d(d+1)}{2} - 1\right) \cdot \binom{n-3d-1}{k-2} + C\frac{k^3}{n^3}\binom{n}{k},$$

then there exists $e \in [n]$ such that the juntas $\mathcal{G}_1, \mathcal{G}_2$ are contained in the same $\mathcal{U}_{e,d,k}$.

Proof of Proposition 4.3.2. It is sufficient to prove the 'moreover' statement. Write $\mathcal{G}_i = \langle \mathcal{J}_i \rangle$ for some family $\mathcal{J}_i \subseteq \mathcal{P}(J)$, and suppose that

$$\min_{i=1,2} |\mathcal{G}_i| \geq \binom{n}{k} - \binom{n-d-1}{k} + \left(\frac{d(d+1)}{2} - 1\right) \binom{n-3d-1}{k-2} + C\frac{k^3}{n^3}\binom{n}{k}.$$

Claim 4.3.4. *Each family $\mathcal{J}_i$ contains at least $d+1$ singletons.*

Proof. Suppose there exists $i \in \{1,2\}$ such that $\mathcal{J}_i$ has at most d singletons, then

$$\begin{aligned}|\mathcal{G}_i| &\leq d \cdot \binom{n-1}{k-1} + \binom{j}{2}\binom{n-2}{k-2} \\ &\leq \binom{n}{k} - \binom{n-d-1}{k},\end{aligned}$$

contradicting the hypothesis. Indeed the final inequality is true by the following reasoning. For sufficiently large C we have

$$\begin{aligned}&\left(\binom{j}{2} + \binom{d}{2}\right)\binom{n-2}{k-2} \leq \binom{n-d-1}{k-1} \\ \Rightarrow &\binom{j}{2}\binom{n-2}{k-2} \leq \binom{n-d-1}{k-1} - (1+2+3+...+(d-1))\binom{n-2}{k-2}. \end{aligned} \tag{4.3}$$

Now, as for each $1 \leq l \leq d-1$ we have

$$l\binom{n-2}{k-2} \geq \binom{n-2}{k-2} + \binom{n-3}{k-2} + ... + \binom{n-l-1}{k-2} = \binom{n-1}{k-1} - \binom{n-l-1}{k-1},$$

the right hand side of inequality (4.3) is easily bounded above by

$$\begin{aligned}\text{RHS} &\leq \binom{n-d-1}{k-1} - \sum_{l=1}^{d-1}\left[\binom{n-1}{k-1} - \binom{n-l-1}{k-1}\right] \\ &= \left[\sum_{l=0}^{d}\binom{n-l-1}{k-1}\right] - d \cdot \binom{n-1}{k-1} \\ &= \left[\binom{n}{k} - \binom{n-d-1}{k}\right] - d \cdot \binom{n-1}{k-1}. \end{aligned} \tag{4.4}$$

Hence, by combining inequalities (4.3) and (4.4) we see

$$\begin{aligned}&\binom{j}{2}\binom{n-2}{k-2} \leq \left[\binom{n}{k} - \binom{n-d-1}{k}\right] - d \cdot \binom{n-1}{k-1} \\ \Rightarrow\ &d \cdot \binom{n-1}{k-1} + \binom{j}{2}\binom{n-2}{k-2} \leq \binom{n}{k} - \binom{n-d-1}{k}\end{aligned}$$

as claimed. □

Now, if $\{a\} \in \mathcal{J}_1, \{b\} \in \mathcal{J}_2$ are singletons, we must have $\text{dist}(a,b) \leq d$, otherwise,

for C sufficiently large, there exist $A, B \in ([n] \setminus D_d(J))^{(k-1)}$ that are not d-close, and so $A \cup \{a\} \in \mathcal{G}_1, B \cup \{b\} \in \mathcal{G}_2$ are not d-close, a contradiction.

Claim 4.3.5. *There exists some e such that the singletons in $\mathcal{J}_1$ and $\mathcal{J}_2$ are*

$$\{e\}, \{e+1\}, ..., \{e+d\}$$

Proof of Claim. Suppose $\{a\} \in \mathcal{J}_1$ and $\{b\} \in \mathcal{J}_2$ such that $\text{dist}(a, b)$ is maximised over singletons in $\mathcal{J}_1$ and $\mathcal{J}_2$ and suppose without loss of generality that this distance is clockwise.

Suppose first that for all singletons $\{a'\} \in \mathcal{J}_1$ and $\{b'\} \in \mathcal{J}_2$ we have $a \leq a', b' \leq b$ (i.e., a is anticlockwise from a', b', which in turn are anticlockwise from b). Then write the first $d+1$ singletons of $\mathcal{J}_1$ in clockwise order starting from $a = a_0 < a_1 < ... < a_d \leq b$. Then $\text{dist}(a, b) \geq d$, and so combining this with the fact that $\text{dist}(a, b) \leq d$ we have equality. We can then see the conclusion must hold with $e = a$ and $e + d = b$.

Suppose instead that there exists $\{a'\} \in \mathcal{J}_1$ such that $a < b < a'$. If there exists singleton $\{b'\} \in \mathcal{J}_2$ such that $b' \leq a < b < a'$ and n is sufficiently large with respect to d then $\text{dist}(a', b') > \text{dist}(a, b)$, a contradiction. Hence for every singleton $\{b'\} \in \mathcal{J}_2$ we have $a < b' \leq b < a'$. Then write the first $d+1$ singletons of $\mathcal{J}_2$ in clockwise order starting from a as $a < b_0 < b_1 < ... < b_d \leq b < a'$, and we see that $\text{dist}(a, b) \geq d + 1$, a contradiction.

The conclusion then follows in general, with $e = a, e + d = b$. □

Claim 4.3.6. *Each family $\mathcal{J}_i$ contains at least $\frac{d(d+1)}{2}$ pairs disjoint from $\{e, e+1, ..., e+d\}$*

Proof. Suppose there exists $i \in \{1, 2\}$ such that $\mathcal{J}_i$ contains fewer than $\frac{d(d+1)}{2}$ pairs

disjoint from $\{e, e+1, ..., e+d\}$. Then

$$\begin{aligned} |\mathcal{G}_i| &\leq \binom{n}{k} - \binom{n-d-1}{k} + \left(\frac{d(d+1)}{2} - 1\right)\binom{n-2}{k-2} + \binom{j}{3}\binom{n-3}{k-3}, \\ &\leq \binom{n}{k} - \binom{n-d-1}{k} + \left(\frac{d(d+1)}{2} - 1\right)\binom{n-3d-1}{k-2} + C\frac{k^3}{n^3}\binom{n}{k}, \end{aligned} \tag{4.5}$$

contradicting the hypothesis when C is sufficiently large. Indeed the final inequality is true if and only if

$$\begin{aligned} \binom{j}{3}\binom{n-3}{k-3} &\leq C\frac{k^3}{n^3}\binom{n}{k} - \left(\frac{d(d+1)}{2} - 1\right)\left(\binom{n-2}{k-2} - \binom{n-3d-1}{k-2}\right) \\ &= C\frac{k^3}{n^3}\binom{n}{k} - \left(\frac{d(d+1)}{2} - 1\right)\left[\sum_{l=3}^{3d+1}\binom{n-l}{k-3}\right]. \end{aligned} \tag{4.6}$$

Now the right hand side of inequality (4.6) is easily lower bounded by

$$C\frac{k^3}{n^3}\binom{n}{k} - \left(\frac{d(d+1)}{2} - 1\right)\cdot(3d-1)\cdot\binom{n-3}{k-3}$$

and since for sufficiently large C we have

$$\left(\binom{j}{3} + \left(\frac{d(d+1)}{2} - 1\right)\cdot(3d-1)\right)\binom{n-3}{k-3} \leq C\frac{k^3}{n^3}\binom{n}{k}$$

we see that inequality (4.6) holds, and thus so does inequality (4.5). This contradicts the hypothesis, and we deduce that $\mathcal{J}_i$ contains at least $\frac{d(d+1)}{2}$ pairs disjoint from $\{e, ..., e+d\}$. □

Now if $\{a, a'\} \in \mathcal{J}_1$ is a pair disjoint from $\{e, ..., e+d\}$ (every element of which is a singleton in $\mathcal{J}_2$), then

$$\{e, e+1, ..., e+d\} \subseteq D_d(\{a, a'\}), \tag{4.7}$$

i.e., $\{a, a'\} \in \mathcal{P}_{e,d}$.

Indeed, if for some $t \in \{0, 1, ..., d\}$ we have $e+t \notin D_d(\{a, a'\})$ then for C sufficiently

large there exist $A \in ([n] \setminus D_d(J))^{(k-2)}$ and $B \in ([n] \setminus D_d(J))^{(k-1)}$ that are not d-close, and so $A \cup \{a, a'\} \in \mathcal{G}_1, B \cup \{e+t\} \in \mathcal{G}_2$ are not d-close, a contradiction.

Hence the pairs in $\mathcal{J}_1$ that are disjoint from $\{e, e+1, ..., e+d\}$ are contained in $\mathcal{P}_{e,d}$. Since there are at least $\frac{d(d+1)}{2} = |\mathcal{P}_{e,d}|$ such pairs, we see that in fact that all pairs in $\mathcal{P}_{e,d}$ lie in $\mathcal{J}_1$. Identical reasoning proves the same result for $\mathcal{J}_2$.

Claim 4.3.7. *Both $\mathcal{G}_1$ and $\mathcal{G}_2$ are contained in $\mathcal{U}_{e,d,k}$.*

Proof. Suppose for a contradiction that $A \in \mathcal{G}_1 \setminus \mathcal{U}_{e,d,k}$, so

1. $A \subseteq [n], |A| = k$,
2. $A \cap \{e, e+1, ..., e+d\} = \emptyset$,
3. For all $\{s,t\}$ such that $\{e, e+1, ..., e+d\} \subseteq D_d(\{s,t\})$ we have $\{s,t\} \not\subset A$.

Suppose $\{e, e+1, ..., e+d\} \subseteq D_d(A)$. Then let $a \in A$ minimise $\text{dist}(a, e)$ and $b \in A$ minimise $\text{dist}(b, e+d)$. Then there exists $0 \leq t \leq d$ such that $e+t \notin D_d(\{a,b\})$. On the other hand, since $e+t \in D_d(A)$, there exists $c \in A$ such that $\text{dist}(c, e+t) \leq d$.

Now we can't have $a < c < e$ as this contradicts the definition of a, we can't have $e \leq c \leq e+d$ as this contradicts the 2nd fact above, and we can't have $e+d < c < b$ as this contradicts the definition of b. It follows that at least one of $\text{dist}(a, e+t)$ or $\text{dist}(b, e+t)$ is at most $\text{dist}(c, e+t) \leq d$, and so $e+t \in D_d(\{a,b\})$, a contradiction.

Hence, $\{e, e+1, ..., e+d\} \not\subseteq D_d(A)$, and therefore there exists $0 \leq t \leq d$ such that $e+t \notin D_d(A)$. But then, for n sufficiently large, we can find $B \in ([n] \setminus J)^{(k-1)}$ that is not d-close to A, from which it follows that $A \in \mathcal{G}_1$ and $B \cup \{e+t\} \in \mathcal{G}_2$ are not d-close. This contradicts the hypothesis that $\mathcal{G}_1$ and $\mathcal{G}_2$ are d-close, and so we deduce $\mathcal{G}_1 \subseteq \mathcal{U}_{e,d,k}$.

Identical reasoning proves that $\mathcal{G}_2 \subseteq \mathcal{U}_{e,d,k}$. □

This completes the proof of the proposition. □

4.3.3 Bootstrapping to an exact result

In the following subsection we bootstrap our results to an exact result. Our approach is to show that if $\mathcal{B}_1, \mathcal{B}_2$ is a cross d-close pair of k-uniform families such that $|\mathcal{B}_1|$ is very large, then $|\mathcal{B}_2|$ must be very small.

In particular, we prove the following proposition.

Proposition 4.3.3. *For each $r \in \mathbb{N}$, there exists a constant $C = C(r)$, such that the following holds. Let $\varepsilon > 0$, let $k_1, k_2 \leq n/C$ and let $\mathcal{B}_1 \subseteq [n]^{(k_1)}$ and $\mathcal{B}_2 \subseteq [n]^{(k_2)}$ be families that are cross d-close. If $\mu(\mathcal{B}_1) \geq 1 - \varepsilon$, then $\mu(\mathcal{B}_2) \leq O_r(\varepsilon^r)$.*

To prove Proposition 4.3.3 we require the following two lemmas. The first is of Keller and Lifshitz [54]; it considers a monotone family $\mathcal{F} \subseteq \mathcal{P}([n])$ and bounds $\mu(\mathcal{F}^{(k)})$ from below in terms of $\mu(\mathcal{F}^{(l)})$ for $l \leq k$. The second lemma, originally proved in a slightly different form by Friedgut [33, 34], gives an upper bound for the size of a k uniform family $\mathcal{F}$ in terms of $\mu_p(\mathcal{F}^{\uparrow})$, where p is approximately k/n.

Lemma 4.3.1. *For any constants $\zeta > 0$ and $r \in \mathbb{N}$, there exists a constant $C(\zeta, r) > 0$, such that the following holds. Let $\zeta n < k \leq (1-\zeta)n$, let $\varepsilon > 0$ and let $l < k/C$. Suppose that $\mathcal{F} \subseteq \mathcal{P}([n])$ is a monotone family that satisfies $\mu(\mathcal{F}^{(k)}) \leq \varepsilon$. Then $\mu(\mathcal{F}^{(l)}) \leq O_{\zeta,r}(\varepsilon^r)$*

Lemma 4.3.2. *Let $n, k \in \mathbb{N}$ and suppose that $0 < p, \phi < 1$ satisfy*

$$p \geq \frac{k}{n} + \frac{\sqrt{2n\log(1/\phi)}}{n}.$$

Then for any family $\mathcal{F} \subseteq [n]^{(k)}$, we have

$$\mu_p(\mathcal{F}^{\uparrow}) > (1-\phi)\frac{|\mathcal{F}|}{\binom{n}{k}} = (1-\phi)\mu(\mathcal{F}).$$

Proof of Proposition 4.3.3. Let $\varepsilon > 0$, let $k_1, k_2 \leq n/C$ and let $\mathcal{B}_1 \subseteq [n]^{(k_1)}$ and $\mathcal{B}_2 \subseteq [n]^{(k_2)}$ be families that are cross d-close. Also, let $p = \frac{1}{2} \cdot \frac{1}{3^{2d+3}+1}$. Consider $\tilde{\mathcal{B}_2} = \left(\mathcal{B}_2^{\uparrow}\right)^{(pn)}$.

By Lemma 4.3.1 we have

$$\mu(\mathcal{B}_2) \leq O_r\left(\mu\left(\tilde{\mathcal{B}_2}\right)^r\right)$$

as long as C is sufficiently large.

Claim 4.3.8. $\mu\left(\tilde{\mathcal{B}_2}\right) = O(\varepsilon)$

Proof. We have $\mu(\mathcal{B}_1) \geq 1-\varepsilon$. Suppose that

$$2p \geq p_1 > \frac{k_1}{n} + \frac{\sqrt{2n\log(1/\varepsilon)}}{n}, \quad 2p \geq p_2 > p + \frac{\sqrt{2n\log(1/\varepsilon)}}{n},$$

then, by Lemma 4.3.2 and since $\mu_{p_1} \preceq \nu_d$ we see firstly that

$$\nu_d(\mathcal{B}_1^{\uparrow}) \geq \mu_{p_1}(\mathcal{B}_1^{\uparrow}) > (1-\varepsilon)\mu(\mathcal{B}_1) \geq (1-\varepsilon)(1-\varepsilon).$$

Secondly, as $\mathcal{B}_1^{\uparrow}$ and $\tilde{\mathcal{B}_2}^{\uparrow}$ are cross d-close we have $\nu_d(\mathcal{B}_1^{\uparrow}) + \nu_d(\tilde{\mathcal{B}_2}^{\uparrow}) \leq 1$, and so $\nu_d(\tilde{\mathcal{B}_2}^{\uparrow}) \leq 1-(1-\varepsilon)(1-\varepsilon)$. Now, by a second application of Lemma 4.3.2 and using $\mu_{p_2} \preceq \nu_d$ we have $\nu_d(\tilde{\mathcal{B}_2}^{\uparrow}) \geq \mu_{p_2}(\tilde{\mathcal{B}_2}^{\uparrow}) > (1-\varepsilon)\mu(\tilde{\mathcal{B}_2})$. Hence

$$\mu(\tilde{\mathcal{B}_2}) < \frac{1-(1-\varepsilon)(1-\varepsilon)}{(1-\varepsilon)} = O(\varepsilon).$$

□

Hence $\mu(\mathcal{B}_2) \leq O_r(\varepsilon^r)$ as claimed, completing the proof of the proposition. □

4.3.4 Families $\mathcal{U}_{e,d,k}$ are locally maximal among d-close familes

We now use the bootstrapping result of the previous subsection to show that when $\mathcal{F}_1, \mathcal{F}_2$ are small alterations of a $\mathcal{U}_{e,d,k}$ (for some $e \in [n]$), then

$$\min_{i\in\{1,2\}} |\mathcal{F}_i| \leq |\mathcal{U}_{e,d,k}|,$$

where equality holds if and only if $\mathcal{F}_1, \mathcal{F}_2 = \mathcal{U}_{e,d,k}$ for some $e \in [n]$.

Proposition 4.3.4. *There exists a constant C such that the following holds. Let $k \leq n/C$ and let $\mathcal{F}_1, \mathcal{F}_2 \subseteq [n]^{(k)}$ be families that are cross d-close. Suppose additionally that there exists a set $U = \{e-d, e-d+1, ..., e+2d\}$ such that for $i = 1, 2$*

$$\mu\left((\mathcal{F}_i)_U^T\right) \leq O\left(\left(\frac{k}{n}\right)^3\right)$$

for all $T \subseteq U$ such that $\{e, e+1, ..., e+d\} \not\subseteq D_d(T)$. Then

$$\min_{i=1,2} |\mathcal{F}_i| \leq \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1}$$

with equality if and only if the families $\mathcal{F}_1, \mathcal{F}_2$ are both equal to the same $\mathcal{U}_{e,k,d}$.

Proof. Let $\mathcal{F}_1, \mathcal{F}_2 \subseteq [n]^{(k)}$ be cross d-close and suppose that

$$\min_{i=1,2} |\mathcal{F}_i| \geq \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1}.$$

Let $U = \{e-d, e-d+1, ..., e+2d\}$ as in the statement. We show that both the families are equal to $\mathcal{U}_{e,d,k}$. Write

$$\max_{i=1,2} \mu\left((\mathcal{F}_i)_U^T\right) = \varepsilon_T$$

for each $T \subseteq U$ such that $\{e, e+1, ..., e+d\} \not\subseteq D_d(T)$. Set

$$V = \{T \subseteq U \text{ such that } \{e, e+1, ..., e+d\} \not\subseteq D_d(T)\}.$$

Now

$$\begin{aligned} |\mathcal{F}_1| + |\mathcal{F}_2| &= \sum_{T \subseteq U} |\,(\mathcal{F}_1)_U^T\,| + |\,(\mathcal{F}_2)_U^T\,| \\ &= \sum_{T \in V} |\,(\mathcal{F}_1)_U^T\,| + |\,(\mathcal{F}_2)_U^T\,| + \sum_{T \notin V} |\,(\mathcal{F}_1)_U^T\,| + |\,(\mathcal{F}_2)_U^T\,|, \end{aligned} \tag{4.8}$$

and we can, by assumption, upper bound

$$\sum_{T\in V} |\,(\mathcal{F}_1)_U^T\,| + |\,(\mathcal{F}_2)_U^T\,| \leq 2\cdot\sum_{T\in V}\varepsilon_T\binom{n-|U|}{k-|T|}. \tag{4.9}$$

In order to bound

$$\sum_{T\notin V} |\,(\mathcal{F}_1)_U^T\,| + |\,(\mathcal{F}_2)_U^T\,| \tag{4.10}$$

we note that for each $S\in V$ either $\mu\left((\mathcal{F}_1)_U^S\right)=\varepsilon_S$ or $\mu\left((\mathcal{F}_2)_U^S\right)=\varepsilon_S$. Let i_S be 1 if $\mu\left((\mathcal{F}_1)_U^S\right)=\varepsilon_S$, otherwise let $i_S=2$. Let $\overline{i_S}$ be the remaining index.

The following pairs of families are cross d-close

$$\left\{\left(\mathcal{F}_{\overline{i_S}}\right)_U^R, (\mathcal{F}_{i_S})_U^S\right\} \;:\; S\in V, \emptyset\neq R\subseteq\{e,e+1,...,e+d\}\setminus D_d(S)$$

so by applying Proposition 4.3.3 with $r=2$ we see that

$$\mu\left(\left(\mathcal{F}_{\overline{i_S}}\right)_U^R\right)\leq 1-\Omega(\sqrt{\varepsilon_S})\Rightarrow |\left(\mathcal{F}_{\overline{i_S}}\right)_U^R|\leq\binom{n-|U|}{k-|R|}\cdot(1-c_S\sqrt{\varepsilon_S})\,.$$

Hence, we can upper bound sum (4.10) by

$$2\cdot\left[\binom{n}{k}-\binom{n-2d-1}{k}-d\cdot\binom{n-2d-2}{k-1}\right]-\sum_{R\in W}\max_{S\in V:D_d(S)\cap R=\emptyset}\binom{n-|U|}{k-|R|}\cdot c_S\sqrt{\varepsilon_S},$$

where $W=\{\emptyset\neq R\subseteq\{e,e+1,...,e+d\}:\exists S\in V$ such that $D_d(S)\cap R=\emptyset\}$. We can combine this bound and the upper bound (4.9) to upper bound the sum (4.8) as follows

$$\begin{aligned}|\mathcal{F}_1|+|\mathcal{F}_2|\leq 2\cdot\sum_{T\in V}\varepsilon_T\binom{n-|U|}{k-|T|}-\sum_{R\in W}\max_{S\in V:D_d(S)\cap R=\emptyset}\binom{n-|U|}{k-|R|}\cdot c_S\sqrt{\varepsilon_S}\\ +2\cdot\left[\binom{n}{k}-\binom{n-2d-1}{k}-d\cdot\binom{n-2d-2}{k-1}\right]\end{aligned} \tag{4.11}$$

Now consider the term

$$2 \cdot \sum_{T \in V} \varepsilon_T \binom{n-|U|}{k-|T|} - \sum_{R \in W} \max_{S \in V : D_d(S) \cap R = \emptyset} \binom{n-|U|}{k-|R|} \cdot c_S \sqrt{\varepsilon_S}. \tag{4.12}$$

We define for each $T \in V$ the sets $W_T := \{R \in W : D_d(T) \cap R = \emptyset\}$. We can then rearrange (4.12) and bound above by

$$\begin{aligned}
&\sum_{T \in V} \left[2 \cdot \varepsilon_T \binom{n-|U|}{k-|T|} - \sum_{R \in W_T} \frac{1}{|\{S \in V : D_d(S) \cap R = \emptyset\}|} \binom{n-|U|}{k-|R|} \cdot c_T \sqrt{\varepsilon_T} \right], \\
&\leq \sum_{T \in V} \left[2 \cdot \varepsilon_T \binom{n-|U|}{k-|T|} - \frac{1}{2^{2d}} \binom{n-|U|}{k-1} \cdot c_T \sqrt{\varepsilon_T} \right],
\end{aligned}$$

where the last inequality uses the fact that $|\{S \in V : D_d(S) \cap R = \emptyset\}| \leq 2^{2d}$ for all R, and that for each $T \in V$ there exists an R of size 1 such that $D_d(T) \cap R = \emptyset$. We now use the fact that $\varepsilon_T = O\left(\left(\frac{k}{n}\right)^3\right)$ to see that each term in the above sum is negative since

$$2 \cdot \varepsilon_T \binom{n-|U|}{k-|T|} - \frac{1}{2^{2d}} \binom{n-|U|}{k-1} \cdot c_T \sqrt{\varepsilon_T} \leq O\left(\left(\frac{k}{n}\right)^3\right) \binom{n-|U|}{k} - \binom{n-|U|}{k} O\left(\left(\frac{k}{n}\right)^{5/2}\right) \leq 0,$$

and equality occurs only if $\varepsilon_T = 0$ for all $T \in V$.

Substituting this back into inequality (4.11) we see that

$$\begin{aligned}
|\mathcal{F}_1| + |\mathcal{F}_2| &\leq 2 \cdot \left[\binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1} \right] \\
\Rightarrow \min_{i=1,2} |\mathcal{F}_i| &\leq \binom{n}{k} - \binom{n-2d-1}{k} - d \cdot \binom{n-2d-2}{k-1},
\end{aligned}$$

with equality if and only if $\varepsilon_T = 0$ for all $T \in V$, which in turn is if and only if $\mathcal{F}_1, \mathcal{F}_2 = \mathcal{U}_{e,k,d}$ for some value of e.

□

4.3.5 Proof of the main theorem

We are now ready to prove the main theorem.

Proof of Theorem 4.1.2. By Theorem 4.3.1 applied with $r = 3$ there exists an $O(1)$-set J and juntas $\mathcal{J}_1, \mathcal{J}_2 \subseteq \mathcal{P}(J)$ that are cross d-close, such that $\mu\left(\mathcal{F}_i \setminus \left\langle \mathcal{J}_i^{\uparrow} \right\rangle\right) = O\left(\left(\frac{k}{n}\right)^3\right)$. Assuming that

$$\min_{i=1,2} |\mathcal{F}_i| \geq \binom{n}{k} - \binom{n-d-1}{k} + \frac{d(d+1)}{2} \cdot \binom{n-3d-1}{k-2}$$

this implies that

$$\begin{aligned}\left|\left\langle \mathcal{J}_i^{\uparrow} \right\rangle\right| &\geq \binom{n}{k} - \binom{n-d-1}{k} + \frac{d(d+1)}{2} \cdot \binom{n-3d-1}{k-2} - O\left(\left(\frac{k}{n}\right)^3 \binom{n}{k}\right) \\ &\geq \binom{n}{k} - \binom{n-d-1}{k} + \left(\frac{d(d+1)}{2} - 1\right) \cdot \binom{n-3d-1}{k-2} + C\frac{k^3}{n^3}\binom{n}{k},\end{aligned}$$

for each $i \in \{1, 2\}$. By Proposition 4.3.2, it follows that the juntas $\left\langle \mathcal{J}_1^{\uparrow} \right\rangle, \left\langle \mathcal{J}_2^{\uparrow} \right\rangle$ are all equal to the same $\mathcal{U}_{e,k,d}$. Provided that C is sufficiently large, the assertion of the theorem follows now from Proposition 4.3.4. □

4.4 Conclusion

We have shown that for each positive integer n and integer d, if k is a non-negative integer that is sufficiently small with respect to n then k-uniform d-close families have size at most that of families of the form $\mathcal{U}_{e,d,k}$. Furthermore, when such families are close to maximum size, then they are small alterations of some particular $\mathcal{U}_{e,d,k}$ (specified by a particular $e \in [n]$). It would be interesting to extend this exact result to all values of $k \in \{0\} \cup [n]$. Clearly, the answer is trivially that for $k \geq \frac{n-1}{2} - (2d-1)$, a k-uniform d-close family has size at most $\binom{n}{k}$ and this is tight as $[n]^{(k)}$ is itself d-close.

4.5 Acknowledgments

The work in this chapter is based on ongoing joint work between David Ellis and the author.

References

[1] James Aaronson, Carla Groenland, and Tom Johnston. Cyclically covering subspaces in $\mathbb{F}_2^n$. *arXiv:1903.10613 [math.CO]*, 2019.

[2] Foto N. Afrati, Anish Das Sarma, Semih Salihoglu, and Jeffrey D. Ullman. Upper and Lower Bounds on the Cost of a Map-Reduce Computation. *Proceedings of the VLDB Endowment*, 6(4):277–288, 2013.

[3] Rudolf Ahlswede and Ning Cai. A Counterexample to Kleitman's Conjecture Concerning an Edge-Isoperimetric Problem. *Combinatorics, Probability and Computing*, 8(04):301–305, 1999.

[4] Rudolf Ahlswede and Ning Cai. Appendix: On Edge-Isoperimetric Theorems for Uniform Hypergraphs. In *General Theory of Information Transfer and Combinatorics*, pages 979–1005. Springer, 2006.

[5] Rudolf Ahlswede and Gyula Katona. Graphs with Maximal Number of Adjacent Pairs of Edges. *Acta Mathematica Hungarica*, 32(1-2):97–120, 1978.

[6] Rudolf Ahlswede and Levon H. Khachatrian. The complete nontrivial-intersection theorem for systems of finite sets. *Journal of Combinatorial Theory, Series A*, 76(1):121–138, 1996.

[7] Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *European Journal of Combinatorics*, 18(2):125–136, 1997.

[8] Ian Anderson. *Combinatorics of Finite Sets*. Courier Corporation, 1987.

[9] Kenneth Arrow. A difficulty in the concept of social welfare. *The Journal of Political Economy*, 58(4):328–346, 1950.

[10] Paul Beame and Cyrus Rashtchian. Massively-Parallel Similarity Join, Edge-

Isoperimetry, and Distance Correlations on the Hypercube. In *Proceedings of the Twenty-Eighth ACM-SIAM Symposium on Discrete Algorithms*, pages 289–306, 2017.

[11] Arthur J. Bernstein. Maximally Connected Arrays on the n-cube. *SIAM Journal of Applied Mathematics*, 15(6):1485–1489, 1967.

[12] Sergei Bezrukov. Edge Isoperimetric Problems on Graphs. *Graph Theory and Combinatorial Biology*, 7:157–197, 1999.

[13] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correction with applications to numerical problems. *Proceedings of the 22nd Annual Association for Computing Machinery Symposium on Theory of Computing*, pages 73–83, 1990.

[14] Tom Bohman and Ryan R. Martin. A Note on G-intersecting Families. *Discrete Mathematics*, 260:183–188, 2003.

[15] Béla Bollobás and Imre Leader. Compressions and Isoperimetric Inequalities. *Journal of Combinatorial Theory, Series A*, 56:47–62, 1991.

[16] Béla Bollobás and Imre Leader. Edge-Isoperimetric Inequalities in the Grid. *Combinatorica*, 11(4):299–314, 1991.

[17] Peter J. Cameron, Péter Frankl, and William M. Kantor. Intersecting families of finite sets and fixed-point-free 2 elements. *European Journal of Combinatorics*, 10:149–160, 1989.

[18] Peter J. Cameron and Cheng Y. Ku. Intersecting families of permutations. *European Journal of Combinatorics*, 24 (7):881–890, 2003.

[19] Fan R. K. Chung, Ronald L. Graham, Péter Frankl, and James B. Shearer. Some Intersection Theorems for Ordered Sets and Graphs. *Journal of Combinbatorial Theory, Series A*, 43:23–37, 1986.

[20] George F. Clements and Bernt Lindström. A Generalization of a Combinatorial Theorem of Macaulay. *Journal of Combinatorial Theory*, 7(3):230–238, 1969.

[21] Shagnik Das, Wenying Gan, and Benny Sudakov. The Minimum Number of Disjoint Pairs in Set Systems and Related Problems. *Combinatorica*, 36(6):623–660, 2016.

[22] Michel Deza and Péter Frankl. On the maximum number of permutations with given maximal or minimal distance. *Journal of Combinatorial Theory, Series A*,

22:352–360, 1977.

[23] Irit Dinur and Ehud Friedgut. Intersecting Families are Essentially Contained in Juntas. *Combinatorics, Probability and Computing*, 18 (1-2):107–122, 2009.

[24] Peter J. Cameron (Ed.). Research problems. In *Proceedings of the 13th British Combinatorial Conference*, volume 125, pages 407–417. 1994.

[25] David Ellis, Yuval Filmus, and Ehud Friedgut. Triangle-intersecting families of graphs. *Journal of the European Mathematical Society*, 14(3):841–885, 2012.

[26] David Ellis, Ehud Firedgut, and Haran Pilpel. Intersecting Families of Permutations. *Journal of the American Mathematical Society*, 24(3):649–682, 2011.

[27] David Ellis and Imre Leader. An isoperimetric inequality for antipodal subsets of the discrete cube. *European Journal of Combinatorics*, 70:149–154, 2018.

[28] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *Quarterly Journal of Mathematics. Oxford Second Series*, 12:313–320, 1961.

[29] Burton Fein, William M. Kantor, and Murray Schacher. Relative brauer groups II. *Journal für die reine und angewandte Mathematik*, 328:39–57, 1981.

[30] Cees M. Fortuin, Pieter W. Kasteleyn, and Jean Ginibre. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics*, 22(2):89–103, 1971.

[31] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[32] Ehud Friedgut. Sharp thresholds of graph properties, and the k-SAT problem (with an appendix by Jean Bourgain). *Journal of the American Mathematical Society*, 12(4):1017–1054, 1999.

[33] Ehud Friedgut. On the measure of intersecting families, uniqueness and stability. *Combinatorica*, 28:503–528, 2008.

[34] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society*, 124:2993–3002, 1996.

[35] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connections to learning and approximation. *Journal of the association for computing*

machinery, 45(4):653–750, 1998.

[36] Oded Goldreich and Leonid Levin. A hard-core predicate for all one-way functions. *Proceedings of the 21st Annual Association for Computing Machinery Symposium on Theory of Computing*, pages 25–32, 1989.

[37] William T. Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11:465–588, 2001.

[38] Lawrence H. Harper. Optimal Assignments of Numbers to Vertices. *Journal of the Society for Industrial and Applied Mathematics*, 12(1):131–135, 1964.

[39] Lawrence H. Harper. Optimal numberings and isoperimetric problems on graphs. *Journal of Combinatorial Theory*, 1:385–393, 1966.

[40] Lawrence H. Harper. On a Problem of Kleitman and West. *Discrete Mathematics*, 93(2):169–182, 1991.

[41] Lawrence H. Harper. *Global Methods for Combinatorial Isoperimetric Problems*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2004.

[42] Sergiu Hart. A Note on the Edges of the n-cube. *Discrete Mathematics*, 14(2):157–163, 1976.

[43] Hamed Hatami. A structure theorem for Boolean functions with small total influences. *Annals of Mathematics*, 176(1):509–533, 2012.

[44] Anthony J. W. Hilton and Eric C. Milner. Some intersection theorems for systems of finite sets. *The Oxford Quarterly Journal of Mathematics*, 18(1):369–384, 1967.

[45] Richard Holley. Remarks on the FKG inequalities. *Communications in Mathematical Physics*, 36 (3):227–231, 1974.

[46] Christopher Hooley. On Artin's conjecture. *Journal für die reine und angewandte Mathematik*, 225:209–220, 1967.

[47] I. Martin Isaacs. Finite group theory. In *Graduate Studies in Mathematics*, volume 92. American Mathematical Society, Providence, Rhode Island, 2006.

[48] John R. Isbell. Homogeneous games. *Math. Student*, 25:123–128, 1957.

[49] John R. Isbell. Homogeneous games, II. *Proceedings of the American Mathematical Society*, 11:159–161, 1960.

[50] Robert E. Jamison. Covering finite fields with cosets of subspaces. *Journal of*

Combinatorial Theory, Series A, 22:253–266, 1977.

[51] Jeff Kahn, Gil Kalai, and Nathan Linial. The Influence of Variables on Boolean Functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, SFCS '88, pages 68–80, Washington, DC, USA, 1988. IEEE Computer Society.

[52] Gil Kalai. A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. *Advances in Applied Mathematics*, 29(3):412–426, 2002.

[53] Peter Keevash and Eoin Long. Stability for vertex isoperimetry in the cube. *Journal of Combinatorial Theory, Series B*, 145:113–144, 2020.

[54] Nathan Keller and Noam Lifshitz. The Junta Method in Extremal Hypergraph Theory and Chvátal's Conjecture. *Electronic Notes in Discrete Mathematics*, 61:711–717, 2017.

[55] Naomi Kirshner and Alex Samorodnitsky. A moment ratio bound for polynomials and some extremal properties of Krawchouk polynomials and Hamming spheres. *arXiv:1909.11929 [math.CO]*, 2019.

[56] Daniel J. Kleitman. Families of Non-disjoint Subsets. *Journal of Combinatorial Theory*, 1(1):153–155, 1966.

[57] Imre Leader. Discrete Isoperimetric Inequalities. In B. Bollobás and F.K.R. Chung, editors, *Probabilistic Combinatorics and its Applications*. AMS, 1991.

[58] Rudolf Lidl and Harald Niederreiter. Finite fields. In *Encyclopaedia of Mathematics*, volume 20. Cambridge University Press, Cambridge, 1997.

[59] John H. Lindsey. Assignment of Numbers to Vertices. *The American Mathematical Monthly*, 71(5):508–516, 1964.

[60] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, Fourier transform and learnability. *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, 40(3):607–620, 1993.

[61] Jiang Luh. On the representation of vector spaces as a finite union of subspaces. *Acta Mathematica Hungarica*, 23:341–342, 1972.

[62] Donald J. Newman. Simple analytic proof of the prime number theorem. *American Mathematical Monthly*, 87 (9):693–696, 1980.

[63] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[64] Ryan O'Donnell and Rocco Servedio. Learning monotone decision trees in polynomial time. *SIAM Journal on Computing*, 37(3):827–844, 2007.

[65] Robert Osserman. The isoperimetric inequality. *Bulletin of the American Mathematical Society*, 84(6):1182–1238, 1978.

[66] Klaus Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 28:104–109, 1953.

[67] Robitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

[68] Atle Selberg. An Elementary Proof of the Prime Number Theorem. *Annals of Mathematics*, 50 (2):305–313, 1949.

[69] Pablo Spiga. p-elements in permutation groups, PhD Thesis. Queen Mary, University of London, Available at https://symomega.files.wordpress.com/2010/02/pablosthesis.pdf, 2004.

[70] Michio Suzuki. A new type of simple groups of finite order. *Proceedings of the National Academy of Science, U.S.A.*, 46:868–870, 1960.

[71] John Talbot. Intersecting Families of Separated Sets. *Journal of the London Mathematical Society*, 68(1):37–51, August 2003.

[72] Leslie Valiant. A theory of the learnable. *Communications of the association for computing machinery*, 27(11):1134–1142, 1984.

[73] Da-Lun Wang and Ping Wang. Discrete isoperimetric problems. *SIAM Journal on Applied Math*, 32:860–870, 1977.